Advanced deconvolution techniques and medical radiography

Adrian Leslie Jannetta

Thesis submitted in partial fulfilment of the requirements of the University of Northumbria at Newcastle for the degree of Doctor of Philosophy.

Thesis undertaken in the School of Computing, Engineering and Information Sciences and in collaboration with the Regional Medical Physics Department at Newcastle General Hospital.

October 2005

Abstract

Medical radiography is a process by which the internal structures of the human body are imaged using a source of x-rays. The images formed are essentially shadowgrams whose size and intensity is dependent on the geometry of the imaging system and the degree to which the structures attenuate x-ray radiation. The images are blurred because the x-ray source has a finite size, and noisy because the x-ray exposure must be kept as low as possible for the safety of the patient but which also limits the number of photons available for image formation. In such noisy environments traditional methods of Fourier deconvolution have limited appeal.

In this research we apply maximum entropy methods (MEM) to some radiological images. We justify the choice of MEM over other deconvolution schemes by processing a selection of artificial images in which the blur and noise mimic the real situation but whose levels are known *a priori*. A hybrid MEM scheme is developed to address the shortcomings of so-called historic MEM in these situations.

We initially consider images from situations in which the model point-spread function is assumed to be three-dimensionally spatially invariant, and which approximates the real situation reasonably well. One technique lends itself well to this investigation: magnification mammography. MEM is offered as a way of breaking some of the conflicting performance requirements of this technique and we explore several new system possibilities with a working MEM system in place. A more complicated blurring function is encountered in linear tomography, which uses opposing movements of the image receptor and x-ray source to generate planar images through an object. Features outside a particular focal plane are smeared to such an extent that detail within the focal plane can be very difficult to detect. With appropriate modification of our MEM technique, processed images show a significant reduction to the blurring outside the focal plane.

Contents

*A summary of this work has been published (Jannetta et al. 2004)

List of Figures

Acknowledgements

I acknowledge receipt of a research studentship from Northumbria University. I am grateful to The Royal Society for award of Research Grant 574006.G503-23863/SM for the purchase of dedicated hardware and image processing software. I would like to thank Leeds Test Objects Ltd for permission to reproduce a section of the layout of their TORMAM phantom in this thesis.

Although a few words do not do justice to their contribution, I would like to express my heartfelt thanks to my PhD supervisors. Firstly to John Jackson for his encouragement, guidance and patience throughout period of my research. John's support, which has often been above and beyond the call of duty, has always been greatly valued. Working on this project has been a lot of fun thanks to his boundless enthusiasm. I'm also very grateful for all the help and advice that I've received from John Kotre whenever I have asked for it. John's enthusiastic approach, knowledge and general input (especially in keeping me right on the medical physics side of things) has been invaluable. The relationship between the student and supervisor is often said to be a crucial ingredient for PhD success so I'm very aware of how lucky I've been. I could not have hoped for a better supervisory team.

I would like to express my thanks and appreciation to several people at the Regional Medical Physics Department at Newcastle General Hospital. In particular, thanks to Ian Birch, Rod Padgett and Kevin Robson for tireless running up and down flights of stairs to get my images and for allowing me to be present at every step of the acquisition process (also, for letting me see the guts of a mammography machine!)

It is a pleasure to thank Stephen Gull of the Astrophysics Group, Cavendish Laboratory, University of Cambridge for several very useful conversations about MEM and MEMSYS.

I am grateful to Jon Riley, Jonathan Edwards and John Eakins of the Institute for Image Data Research at Northumbria University for being on hand to answer my queries about image processing issues.

I'd like to say a huge thankyous to my mam and dad, Jen and Les, for supporting me in many ways during the course of my PhD, and to my uncle Norman for giving me the opportunity to get a taste for doing serious research a few years ago.

Finally, I dedicate this work to Emma, my partner and soulmate. Emma has given me inspiration and love when I needed it, and kept me smiling and sane when the going got tough.

Author's declaration

I hereby declare that this thesis, submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy and entitled "Advanced deconvolution techniques and medical radiography", represents my own work and has not been previously submitted to this or any other institution for any degree, diploma or other qualification.

Adrian Leslie Jannetta
October 2005

Acronyms

AEC Automatic exposure control.

BF Broad focal spot setting.

CNR Contrast to noise ratio.

CR Computed Radiography.

CT Computed Tomography.

DFT Discrete Fourier transform.

DQE Detective quantum efficiency.

FF Fine focal spot setting.

FS Focal spot.

FFT Fast Fourier transform.

HFE High frequency emphasis.

ICF Intrinsic correlation function.

MEM Maximum Entropy Method.

MTF Modulation transfer function.

PDF Probability distribution function.

PSF Point-spread function.

SNR Signal to noise ratio.

SVD Singular value deconvolution.

CHAPTER 1

Introduction and research intent

1.1 Research intent

Medical radiography is the process by which the internal structures of the human body can be imaged using a source of x-ray radiation. Structures within the body attenuate x-rays to varying degrees, and the image, recorded using film or a digital receptor, essentially shows shadows whose size and intensity are dependent on the composition of the structures within the body and the characteristics of the imaging system.

Whilst radiography is not a tool used exclusively in a medical context, the quality and properties of the images obtained are determined by both technical considerations and patient-safety concerns which may not be issues in other applications.

A potential remedy to the problems of geometrical blurring and image noise, applicable to modern digital receptors, is digital image processing. Provided the characteristics of the image blurring function are known, an image can, in theory, be restored to that which would have been produced by a perfect point focal spot. As will be seen the presence of image noise complicates the restoration process. Radiological images have a relatively poor signal-to-noise ratio, the reason for which is the requirement to operate within the ALARP principle, so that patient radiation doses are 'kept as low as reasonably practicable consistent with the intended purpose' (IRMER 2000). Radiological images are therefore expected to be inherently noisy, with quantum noise being the dominant noise source.

In this research we apply an image deconvolution technique called max-

imum entropy method (MEM) to some radiological images. We justify our choice of MEM in Chapter 2, by comparing restorations with those obtained with other deconvolution schemes using a selection of artificially degraded images (where the blur and noise levels are known *a priori*). The theoretical derivation and justification of so-called Historic MEM, from an information theory standpoint, is given in Chapter 3.

In Chapter 4 we initially apply MEM to images obtained from situations where the model PSF is assumed to be spatially invariant in three dimensions and which approximates the real situation reasonably well. One technique which lends itself well to this investigation is magnification mammography: this is essentially projection radiography employing sufficient geometric magnification to render fine detail visible on the image receptor. We experiment with the premise of using MEM to break out of the conflicting performance requirements of magnification mammography, and we explore several new system possibilities with a working MEM system in place. With conventional focal spot and magnification settings, we investigate the effect of MEM on image quality at the standard dose and its effect on quality at reduced radiation doses (when quantum noise would otherwise reduce feature visibility). Use of the unconventional broad focal spot setting would allow reduced exposure times if MEM could decrease the greater associated blurring. Similarly, we use MEM to tackle the blurring created by higher geometric magnifications and the loss of image quality associated with radiation dose reduction.

A more complicated blurring function is encountered in tomography. Linear tomography uses opposing movements of the image receptor and x-ray source to generate planar images through an object. Features outside a particular focal plane are smeared out to such an extent that detail within the focal plane can be very difficult to detect. The blurring function in this case is depth dependent and in Chapter 5 we develop an model of the situation and modify the MEM technique appropriately.

The effectiveness of MEM restorations are evaluated through the use of clinical test objects and, in Chapter 6, image data from real radiological applications.

In the rest of this chapter we present some basic principles of radiography and radiological imaging, and focus on the two radiological applications to be examined in detail later in this research: magnification mammography and linear tomography.

1.2 X-ray radiation

X-rays were discovered by Wilhelm Röntgen in 1895 during experiments investigating the effect of cathode rays* that were produced by electrical discharge through gases at low pressure in highly evacuated glass tubes. He discovered that these high energy cathode rays, on striking a material, produced an invisible 'x-ray' which caused fluorescence in a barium platinocyanide coated screen on the other side of the room. Röntgen found that x-rays were electrically neutral and that the rays could penetrate several materials to some extent. Röntgen also observed the fogging effect the rays had on photographic film and saw the shadowed outline of the bones in his hand as he held objects between the x-ray tube and screen. Röntgen published his findings in 1896 and potential applications to medicine were seen immediately. However, problems arising from the dangerous ionizing properties of x-ray radiation were not properly dealt with for some time. The beginnings of x-ray diagnosis are reviewed by Mould (1995).

There are no well defined upper or lower limits in terms of energy or wavelength between which the label 'x-ray' can be applied to a photon. Gamma rays were discovered soon after x-rays. Gamma rays are electromagnetic radiation emitted from the nuclei of some radioactive isotopes. Early experiments showed gamma rays to have higher energies than x-rays. The definition based on energy (and therefore wavelength or frequency) is now somewhat blurred, with some x-ray applications involving higher energies than other experiments using gamma ray radiation.

X-ray production and spectra

X-rays are produced when high energy electrons lose energy during their interaction with atoms in an object. X-ray production occurs when the interaction is one of the following processes.

Deceleration of a fast moving electron The incident electron passes close to the atomic nucleus and the force of attraction causes a deceleration of the electron. The loss of kinetic energy is manifested as an emitted x-ray photon whose energy depends on the size of the deceleration and the attractive force. X-rays produced this way are called Bremsstrahlung ("Braking radiation").

Ionization by removal of K-shell electron The incoming electron interacts with an electron in the K-shell of the atom and transfers energy

*Cathode rays were electrons produced by the negative electrode - the cathode - in a vacuum tube.

to it. If the transferred energy is equal to or greater than the binding energy for an electron in that shell then the electron is ejected and the atom becomes ionised. An electron from an outer shell can fall into the vacancy in the K-shell. The corresponding energy loss from the falling electron is carried away from the atom as a x-ray photon.

Both of these processes take place in the target of an x-ray tube and result in x-ray photons having a range of energies (described by an energy spectrum). Electrons can also interact with atoms in such a way that heat is produced rather that x-ray photons (see section 1.3.1).

X-ray spectra are a superposition of spectra from these two processes. Bremsstrahlung gives rise to a continuous spectrum with photons able to take on energy up to some maximum value which is dependent on the maximum operating voltage of the x-ray tube. The ionization involving an electron from the K-shell gives rise to a type of spectrum which is characteristic of the binding energy within the element from which the target is made. The characteristic spectrum appears as peaks superimposed on the continuous spectrum.

X-ray intensity and quality

The *intensity* of the x-ray beam is a measure of the energy in x-ray photons flowing through a unit area over unit time. The intensity of an x-ray beam is no different from any other electromagnetic radiation in that it follows an inverse-square law as the distance from the focal spot is increased. The beam intensity is also affected by the operating voltage of the x-ray tube and the tube current.

The *quality* of the beam determines how penetrating the beam is. Quality is described by the spectrum of the beam, i.e. in terms of the relative intensities present at each wavelength. Beam quality is affected by the operating voltage. The beam quality can also be modified by filtration techniques. This is usually accomplished by introducing metal foils into the path of the beam (see pages 25 and 177)[†].

1.3 Principles of radiography

In this section some of the basic principles, terminology and problems associated with medical radiography will be introduced. This is not a complete

[†]There is also a small but unavoidable amount of filtration when the x-rays lose energy through interactions with the glass, cooling oil and housing of the x-ray tube.

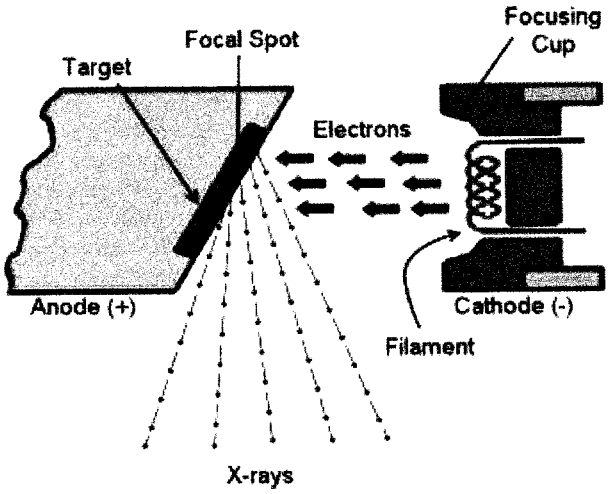

Figure 1.1: Simplified layout of part of an x-ray tube.

introduction to the subject and definitions and concepts discussed in this section relate to the subsequent work. More complete descriptions of the physics and technology of radiography and radiology exist elsewhere; see for example Meredith and Massey (1977); Hay and Hughes (1978); Wilks (1987); Webb (1988); Farr and Allisy-Roberts (1998).

1.3.1 X-ray tube design

Medical x-rays are typically produced in x-ray tubes (see figure 1.1). The tube consists of a filament (cathode), which is heated by an electric current until electrons are released through a process called thermionic emission. Tungsten is often used because it has considerable thermionic emission temperatures well below its melting point. The electrons are accelerated to high energies by the positive potential difference applied to the target (anode). A focussing cup mounted near the filament protects parts of the tube wall from damage by electron bombardment and also directs the electrons with more accuracy towards a small area on the target. The electrons interact with the target in a number of ways. It is from this area - referred to as the *focal spot* - that the x-ray photons emerge. The tube is also a high vacuum to minimise production of electrons by ionization of gas.

Operating voltage

The potential difference between the anode and cathode is typically measured in kilovolts (1kV = 10^3V). The operating voltage may vary with the frequency of the electrical supply so voltages are expressed in terms of the peak voltage and kVp is used to denote this. The peak voltage determines the maximum energy of the emitted x-ray photons. For example, a 50kVp operating voltage can yield photons with energies up to a maximum of 50keV, although the average photon energy is much less than this value. Some simulated x-ray spectra using typical mammography settings can be seen on page 17.

The operating voltage determines the energy spectrum of the x-ray beam. This in turn will affect the radiographic contrast of images obtained from an x-ray unit. If the majority of photons have so little energy that they are absorbed by the object then an image will not be formed. If the energy of the x-ray beam is so high that the intervening material of the object has minimal effect in terms of attenuating the beam then the contrast of the resulting image will be very poor.

Tube current and exposure rate

The electric current passing through the tube is measured in units of milliamperes (mA). For a machine operated at a fixed potential the radiation output rate is proportional to the tube current. The total radiation output per exposure period is measured in units of mAs (milliampere-seconds) and is proportional to the radiation dose received by the object being imaged. The unit quantifying the dose of radiation energy absorbed by an object is the gray (Gy), where 1 Gy = 1 J kg^{-1}.

The radiation dose is determined by the number of photons and their energy spectrum. A brief analysis of the relationship between dose and noise is conducted in section 1.3.3.

Efficiency of the x-ray tube

Electrons can interact with matter in a number of ways. The two processes which lead to x-ray production were described earlier (see page 3). Other interactions between the incident electron and an atom in the target do not lead to x-ray production. They are:

Excitation involving outer shell electrons The incident electron may transfer a small amount of its energy to an electron in an outer shell of the atom. This causes a temporary jump by the orbiting electron

to a level further out from the nucleus. When the electron returns to the lower shell the excess energy is emitted as a photon whose energy is eventually converted to heat by absorption in the target.

Ionisation by outer shell electron removal The incident electron may transfer a small but sufficient quantity of energy to an outer shell electron to ionise the atom. The displaced electron may further interact with other atoms in the target. The small amount of energy lost by the incident electron will eventually contribute heat energy to the system.

The energy carried by the electrons from the filament to the target is converted into x-ray photons and heat. In the diagnostic energy range (≈20 - 150 keV) the above processes account for most of the interactions between the electrons and atoms in the target. The efficiency of an x-ray machine is therefore rather poor, with approximately 99.5% of the total kinetic energy of the electrons being converted to heat and only 0.5% of the energy ending up as x-ray photons due to the processes discussed earlier (page 3).

1.3.2 Geometric magnification and image unsharpness

It is obvious from figure 1.2 that the size of the shadow depends on the position of the object between the x-ray source (the *focal spot*) and image receptor. Moving the object closer to the focal spot will result in a larger, magnified shadow on the image receptor. In the case of the ideal focal spot it can be shown (using similar triangles) that the magnification M is given by:

$$M = \frac{x_i}{x_o} = \frac{a+b}{a} \tag{1.1}$$

In contact radiography the object is close to the image receptor, so $b = 0$, giving a 1:1 image[‡].

In radiography, much of the blurring occurs because the x-ray focal spot is not a point source. The finite size of the focal spot gives features in the image a penumbra. The blurring effects of an actual focal spot are compared to an ideal focal spot in figure 1.2. The degree of blurring varies with the geometry of the focal spot, object and image receptor. Such burring is referred to as *geometric unsharpness*. For a focal spot of fixed size, the amount of blurring is minimised by placing the object as close as possible to the image receptor i.e. contact radiography.

[‡]In reality we will always observe $M > 1$ as b is never reduced to zero: even in contact radiography an air gap (of a number of cms) or, more commonly, an anti-scatter grid is placed between the patient and image receptor.

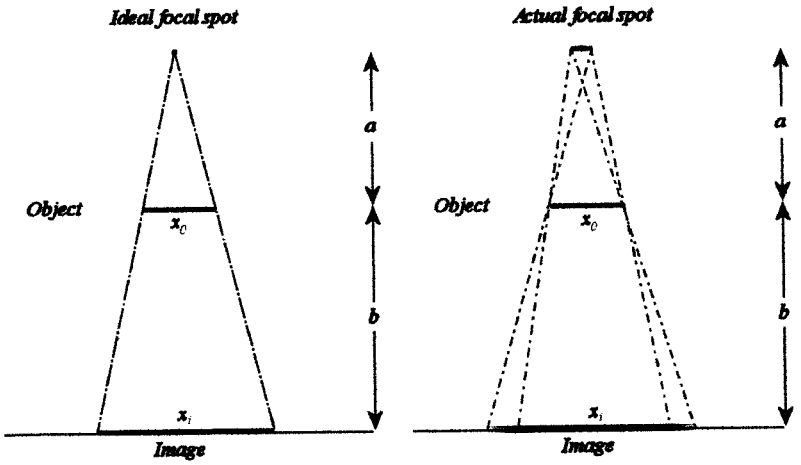

Figure 1.2: Comparison of the effects of using ideal and actual focal spots in radiography.

Other potential sources of image blurring are *movement unsharpness* and *receptor unsharpness*. The former is due the movement of any of the components in figure 1.2 during the exposure. Controlled movement of the focal spot and image receptor is necessary for tomography to work. However, movement unsharpness arising from involuntary movements of the patient during the exposure are not desirable. Movement unsharpness may be minimised by reducing the x-ray exposure time and by taking steps to prevent patient movement. Receptor unsharpness in Computed Radiography arises because the image forming effects of individual x-ray photons are 'amplified' by the receptor to produce a better image. The causes of receptor unsharpness in CR are described in section 1.3.6.

The focal spot

The design of x-ray tubes has evolved to cope with a number of conflicting performance requirements. The focal spot size needs to be as small as possible to reduce geometric unsharpness (see section 1.3.2). However, a small focal spot means the large quantities of heat generated are poorly dissipated in the tube. Low currents, which lead to long exposures, must be used. The most common resulting design solution is a rotating target with an acute target angle. Many x-ray units allow the size of the focal spot to be changed; fine or broad focus may be selected. For example, in the analysis of magnified mammography images in Chapter 4, a laboratory-based Siemens Mammomat

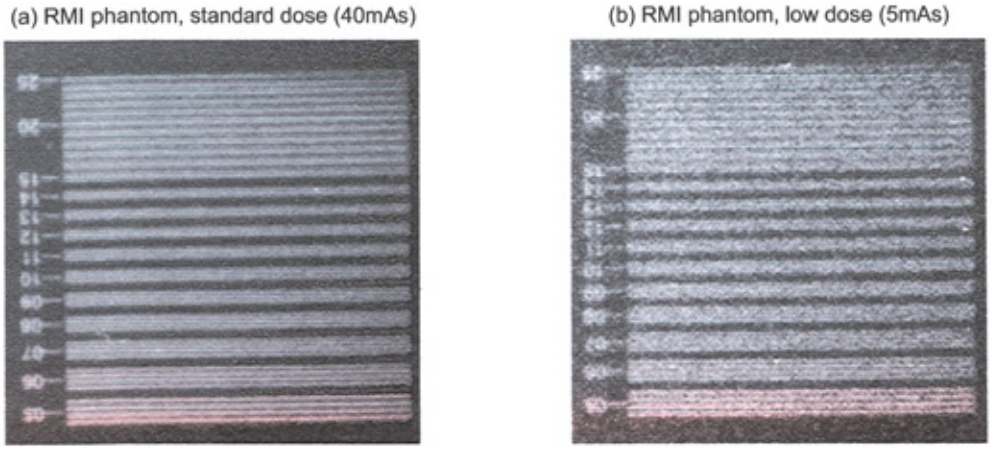

Figure 1.3: A region of the RMI test object imaged using (a) the standard radiation dose and (b) a low radiation dose.

3 mammography unit, with measured focal spot sizes of 0.7×0.3mm (broad focus) and 0.2×0.2mm (fine focus) was used.

1.3.3 Sources of noise

The presence of noise in radiographic images, which manifests itself as graininess or mottling, causes a loss of low contrast detectability. The chief sources of noise in radiographic images are system noise and quantum noise. System noise arises during transmission or digitisation of the image. Quantum noise, also called Poisson noise, is a fine-scale graininess which reflects the fact that we are registering individual x-ray photons in these low intensity beams. The dominant noise source of a good radiological imaging system should be quantum noise. The relationship between radiation dose and quantum noise is explored further in section 1.4.1.

1.3.4 Absorption and scatter processes

When the x-ray beam interacts with atoms in the object being imaged, some x-ray photons will be absorbed by the object's atoms and others will be scattered out of the beam in various directions.

Photoelectric absorption is responsible for most of the object contrast in the resulting image. An x-ray photon is completely absorbed by an atom in the object and an electron is ejected as a result. The energy of the ejected

electron is equal to energy of the original x-ray photon minus the binding energy of the electron to the atom. The binding energy thus defines a threshold, below which there is no photoelectric absorption. The probability of such absorption is greatest just above this threshold, and thereafter decreases with increasing photon energy, and also with increasing atomic number of the scattering atom.

Scattered radiation acts to degrade the image by reducing the contrast. The processes which can scatter the incoming x-ray photons are:

Compton scattering An x-ray photon encountering an atom collides with an electron in the outer shell. The electron bound only loosely to the atom and almost all of the energy transferred to it goes into kinetic energy and the electron is ejected from the atom. The transfer of energy causes a change in energy and direction to the original x-ray photon, both of which depend of the amount of energy transferred (Compton 1923). This is also called incoherent scattering.

Thomson scattering The x-ray photon is completely absorbed by the atom and then re-emitted by the electrons of the atom. There is no net loss of energy so the emitted x-ray photon has the same energy as the original but the direction it follows is random (called elastic scattering). The probability of coherent scattering increases as the incident photon energy decreases and with increasing atomic number of the scattering atom. This is also known as Rayleigh, or coherent scattering.

The relative contributions of each to the total scattered radiation depends on the energy of the x-rays used for a particular application. In diagnostic radiology, with x-ray energies of the order of 100 keV, Compton scattering is predominant with coherent scattering events contributing only a small fraction of the total.

Scatter reduction: grids and air-gaps

Anti-scatter grids are designed to reduce the effects of scattered radiation on image quality. The simplest type of grid is the linear grid[§], which is placed between the object and the image receptor. A linear grid consists of a plate with very thin, highly attenuating, parallel strips of material (usually lead). The space between the strips is filled with a low attenuating material (usually carbon fibre in modern grids). The effectiveness of a grid at removing scattered radiation is quantified by the grid ratio R_{grid} where

[§]Referred to as Bucky grids (Gustave Bucky 1913).

R_{grid} = (strip height)/(strip separation). The height of the strips is always greater than the distance between them and so the grid ratio is always greater than unity. The greater the grid ratio, the better the grid is at reducing scattered radiation.

A variation on the linear grid is the crossed-grid which consists of two linear grids placed perpendicular to one another. Grids containing parallel strips are focussed at infinite distance from the focal spot, i.e. attenuation of non-scattered radiation will occur in most practical situations that they are used. Modern grids are focussed grids, meaning that the lead strips are angled in such a way that they will not attenuate non-scattered radiation from the focal spot. Consequently focussed grids can only be used at predetermined distances from the focal spot.

Images obtained using grids can show the pattern of the grid within the image. The use of moving grids[¶], which shift in position during the radiographic exposure will blur the grid pattern and remove it from visual consideration in captured images.

Grids are used to reduce the effects of scatter in mammography. Despite claims (Chakraborty 1999) that grid removal can improve visibility of small and low contrast features under certain conditions, a more recent experimental comparison of images of the TORMAM test object (see page 108) obtained using a Computed Radiography system, both with and without a grid in place, is described by Coleman et al. (2000). They found that grid-removal had an adverse effect on feature visibility that the corresponding decrease in quantum noise (due to increased dose to the image receptor) did not compensate for.

The presence of an air gap of between 10cm - 20cm between the object and the image receptor will reduce the effects of scattered radiation on image quality. Scattered low energy x-ray photons are allowed to escape from the vicinity of the image receptor without interaction. The increased object-to-receptor distance causes magnification and geometric unsharpness in the image. To compensate, the focal spot to image distance can also be increased. The air-gap technique has been widely used for chest radiography. A model describing the contribution of an air gap on scatter reduction and of the effect of the air gap size is given by Krol et al. (1996).

1.3.5 Subject contrast

X-rays are attenuated by the objects they pass through. The physical processes contributing to this attenuation such as absorption and scatter have been dis-

[¶]Referred to as Hollis-Bucky grids (Hollis Potter 1920).

cussed. Differences in the subject contrast across an object is due to one or more of the following factors:

- Density differences
- Thickness differences
- Atomic number differences
- X-ray quality

Subject contrast, along with receptor contrast, combine to give radiographic contrast. We will discuss contrast further in the following sections.

1.3.6 The image receptor

The first image receptors for use with medical x-rays were direct exposure films. Photographic film offered the advantage of high resolution but required the patient to be subjected to high radiation exposures because the film emulsion is not particularly sensitive to x-ray photons[‖] (Barrett and Swindell 1981). The modern alternatives to film are screen-film, computed radiography and digital image receptors. We describe briefly each of these systems.

Screen-film image receptors

Screen-film systems make use of an fluorescent phosphor *intensifying* screen alongside the photographic film to produce images. The 'heavy' composition of phosphor in the intensifying screen makes it a more efficient absorber of x-ray photons. A single x-ray photon absorbed by the phosphor material can radiate many photons of a lower energy[**] which may interact with a photographic film. Screen-film receptors are actually about 30-50 times more sensitive than standalone photographic film, with immediate advantages being that the radiation dose to the patient can be substantially reduced in order to produce a good quality image and image contrast can be improved by using a lower operating voltage.

The compromise made by using a screen-film system is that an additional source of blurring is introduced into the image. A single x-ray photon may have interacted with a single grain on the film but the introduction of an intensifying screen means that one x-ray photon can potentially affect lots

[‖]A typical emulsion registers around 1% of incident photons at energies of 100keV

[**]Photographic film emulsion is more sensitive to these low energy photons.

of neighbouring grains on the film through emitted photons from the screen. Therefore the screen-film system has an associated point-spread function. This is effect is called receptor unsharpness.

There are other disadvantages associated with film. Film is not linearly sensitive to photon flux; there is a narrow range over which small differences in contrast can be detected. Optimal imaging of tissues with varying densities is difficult to achieve. Practical requirements related to processing time and storage requirements are also issues with screen-film image receptors.

Computed Radiography

Computed radiography (CR) is a process for obtaining digital radiographic images and has been in use since the 1980s. In CR a plate containing photostimulable phosphor (storage phosphor plate) is used instead of film. Unlike the intensifying screens used in the screen-film systems the storage phosphor plate does not emit light when x-rays interact with it, but instead stores the energy of the x-ray photons in a latent image of distributed electron charges. A CR reader uses a laser to release the energy stored by the phosphor plate to create a digital image. This process also introduces a form of blurring into the digital image, due to scattering of the laser beam and the emitted luminescence, which we refer to as receptor unsharpness.

Digitisation of images effectively divides the true distribution (taken to be continuous) into 'bins' or pixels. Pixels are distributed uniformly over the receptor. Using more pixels per unit area leads to a finer resolution and a closer approximation to the continuous distribution of the real object. Each pixel in the CR image has an associated luminance value which is representative of the x-ray intensity averaged over the area of the pixel. The luminance is quantised over a finite number of levels. The level of quantization is usually 2^n where n is the number of *bits* to which the image is digitised. Typically n will take values of 10, 12 or 14 in medical imaging (with $n = 10$ being typical for the image receptors in this work).

For CR to be effective the image receptor (the storage phosphor plate and CR reader taken together) must have good enough spatial resolution for its intended application. Features in the object whose size on the image receptor is less than the receptor resolution will be lost. Similarly, contrast differences in features which are smaller than the intensity quantization resolution will also be lost in the digitisation process.

The pixel values of the image will usually not have a linear relationship to the number of detected x-ray photons. For example, in the case of the Philips ACR-3 CR which was used to obtain many of the images in this work, the relationship is nonlinear and given by:

$$I_{\log} = \left(\frac{1024}{L}\right) \log I_{\text{lin}} - \log\left(\frac{s}{200}\right) + 511 \tag{1.2}$$

where $I_{\log}$ are pixel intensity values of the image. I_{lin} are linear pixel values (related linearly to the x-ray exposure) and s is a measure of the receptor sensitivity. The quantity L is the exposure latitude: the range of film exposure where an increase of log(exposure) will produce an approximately linear increase in film density. The values of L and s are unique to a particular image and are obtained by analysing the histogram of the storage phosphor system. See Samei et al. (2001) for further discussion of these parameters and of image receptor characteristic curves.

When the CR reader reads the plate it is erased and can immediately be used for a capture a new image; there are obvious financial benefits to reusable image receptors. Other advantages of CR arise because storage phosphor plates react to wide range of x-ray energies. Under or over exposures are less critical because the digital image can still be displayed correctly - CR has a wider dynamic range than film. The disadvantages are primarily to do with resolution of the final image, which is not considered good enough for some radiological applications. Yaffe and Rowlands (1997) review the technology and requirements of x-ray detectors. For a review of the physics of CR see Rowlands (2002).

Several types of cassette were used to obtain images with CR in this work. The physical sizes of the cassettes and dimensions of the matrix elements comprising the imaging area of each receptor are listed below:

Type A Physical size: 35cm×35cm. Image dimensions (pixels): 1760×1760. Receptor pixel size = 199μm. This receptor was used in some of our experiments with linear tomography in chapters 5 and 6. The resolution is too low for use in our mammography experiments.

Type B Physical size: 24cm×30cm. Image dimensions (pixels): 1576×1976. Receptor pixel size = 152μm. This receptor was used in some of our experiments involving radiographic imaging of simple test objects in chapter 4 and the linear tomography images of a realistic test object in chapter 6.

Type C Physical size: 18cm×24cm. Image dimensions (pixels): 1770×2370. Receptor pixel size = 101μm. This receptor was used in our mammographic imaging of a clinical test object in chapter 4.

Type D Physical size: 18cm×24cm. Image dimensions (pixels): 3540×4740. Receptor pixel size = 50μm. This receptor was used to obtain the mammograms presented in chapter 6.

Digital receptors

Digital image receptors are used to capture and digitise an image directly during the radiographic exposure. Considerable research has taken place in recent years to the development of technology capable of replacing conventional screen-film systems. We will discuss briefly *indirect* and *direct* digital receptors.

Indirect detectors use a two stage process to obtain an image. An amorphous silicon plate is overlaid with an array of photodiodes, which is itself covered with a layer of CsI(Tl) phosphor. In the first stage the acquisition of x-ray photons will stimulate the emission of light from the phosphor, which is converted to a charge stored in the photodiode. The plate is processed by a reader with each photodiode providing the information for a single pixel in the final image. Direct detectors utilise an amorphous Selenium-coated TFT (thin-film transistor) array to acquire x-ray photons directly and convert them to electric charges. The charges are collected by capacitors linked to each array element and are readout with specialised electronics within the array.

The advantages of digital image receptors are numerous and obvious: images can be captured and displayed very quickly. Radiographic factors such as exposure are less crucial because digital manipulation can compensate for under/over exposure. Problems associated with long term storage of film (physical storage, deterioration of image quality, etc) are also avoided.

1.4 Imaging with CR

1.4.1 Inherent loss of image quality in CR

Several sources of degradation have been considered so far. Most are inherent sources that the radiographer can try to minimise but will be unable to remove completely: geometric unsharpness, motion unsharpness, image receptor blur and x-ray scatter all act to reduce image quality.

Mathematically, noise is defined as a random error of a variable around some mean value. In the context of CR imaging, noise is an unwelcome effect which acts to degrade the image quality. Noise is introduced at every step of the acquisition process and a good CR system should be designed to minimise the effects of noise at each stage.

Quantum noise and Signal-to-noise ratio

The properties of quantum noise, which arise from the random fluctuation of photons counted across the image receptor, is described using Poisson statistics. The standard deviation (σ) from the mean can be used as a measure of the noise. In such a situation the standard deviation from the mean number of photons counted N is given by:

$$\sigma = \sqrt{N} \tag{1.3}$$

So the quantum noise will increase as the square-root of the number of photons increases. The signal, considered to be the difference in detected photons between two regions with a contrast difference, increases as the number of photons increases:

$$\text{Signal} \propto N \tag{1.4}$$

Combining equations (1.3) and (1.4) it can be seen that the signal-to-noise ratio (SNR) varies as:

$$\text{SNR} \propto \frac{N}{\sigma} \propto \frac{N}{\sqrt{N}} \propto \sqrt{N} \tag{1.5}$$

To reduce the effects of quantum noise it is necessary to increase the number of photons from the x-ray source. By equation (1.5), increasing the radiation exposure by a factor of four will double the SNR. Conversely, lowering the radiation exposure by a factor of four will reduce the SNR by a half.

The effect of quantum noise on image quality is seen in figure 1.3, which shows part of the RMI test object imaged using a standard dose and a very low dose. Both images show the inherent graininess associated with Poisson noise but the quality of the image obtained with the low dose is much worse than with the standard dose.

Gaussian approximation

A Poisson distribution describes the statistical properties of quantum noise. If the numbers of photons impinging the image receptor are sufficiently large then we can approximate with a Gaussian distribution. The energy spectrum of a mammography unit operating at 28kVp with a Mo-Mo tube was generated using spectrum processing software provided with IPEM Report 78 (Cranley et al. 1997).

Figure 1.4 shows a plot of photon numbers per 0.5keV bin recorded at the image receptor (per mm^2) against photon energy for two exposures. In

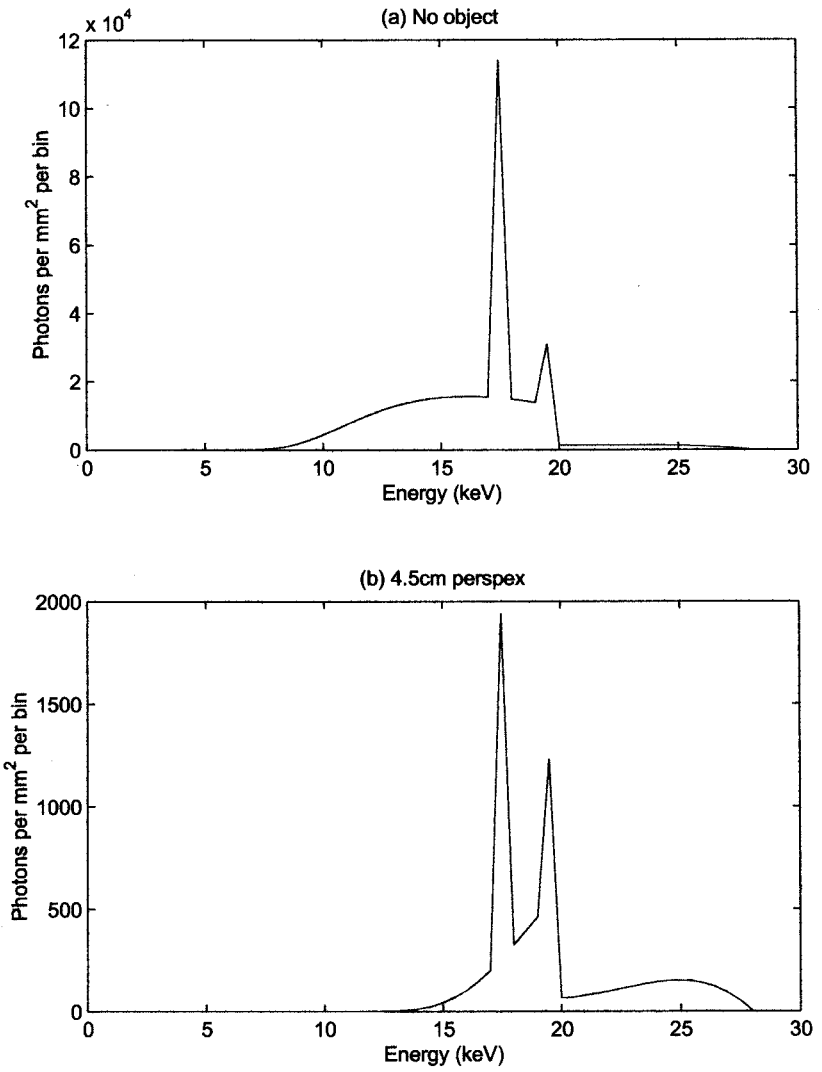

Figure 1.4: X-ray spectra for a mammography unit operating at 28kVp with a Mo-Mo tube. The photon counts derive from bins of 0.5keV. (a) the spectrum produced with no intervening object between the focal spot and image receptor. (b) the spectra obtained with an intervening 4.5cm stack of perspex (equivalent to a standard dosimetry phantom).

the first exposure (figure 1.4a) we have no intervening object between the focal spot and image receptor. The total number of photons in this case is $N_{tot} \approx 620,000$ per mm^2 per bin. In the second exposure (figure 1.4b) the photon counts are those which would be observed if a stack of perspex 4.5cm thick - equivalent to a standard dosimetry phantom - was positioned between the focal spot and image receptor. The total number of photons recorded in this case is $N_{tot} \approx 10,500$ per mm^2 per bin. The corresponding numbers of photons per image pixel (for 100μm receptors) are approximately 6,200 and 105 respectively. These can both be considered 'large' in the context of discussing if Gaussian approximation to a Poisson distribution is appropriate. The latter figure gives a Poisson variation of $\pm 20\%$ (95% confidence), which accounts for the observed granularity in radiological images.

1.4.2 Measures of image quality

There are several ways to quantify image quality; we discuss some these measures below.

MTF and spatial resolution

The modulation transfer function (MTF) is the traditional measure of assessing the spatial resolution of the imaging system. The MTF is defined as the magnitude of the Fourier transform of the point-spread function (PSF)[††]. The MTF describes the ability of an imaging system to reproduce the spatial resolution of the object being imaged. Ideally, the MTF would be constant over all frequencies but in reality it is observed that the magnitude falls off rapidly towards the highest frequencies (see figure 1.5). This behaviour corresponds to a loss of spatial resolution at the finest scales. Factors affecting the MTF and various methods of measuring the MTF are described in Report 41 of the ICRU (1986).

Signal-to-noise ratio

The relationship between SNR and radiation dose was described in section 1.4.1. In practice it is usually difficult to quantify the SNR. The problem is identifying regions of an image containing only signal or noise in order to evaluate the SNR. A compromise is usually to seek a region of the image with constant luminance (signal) and measure the standard deviation (σ) of pixel

[††]The PSF is the imaging system's response to an impulse. The PSF and Fourier transform are discussed in more detail in chapter 2

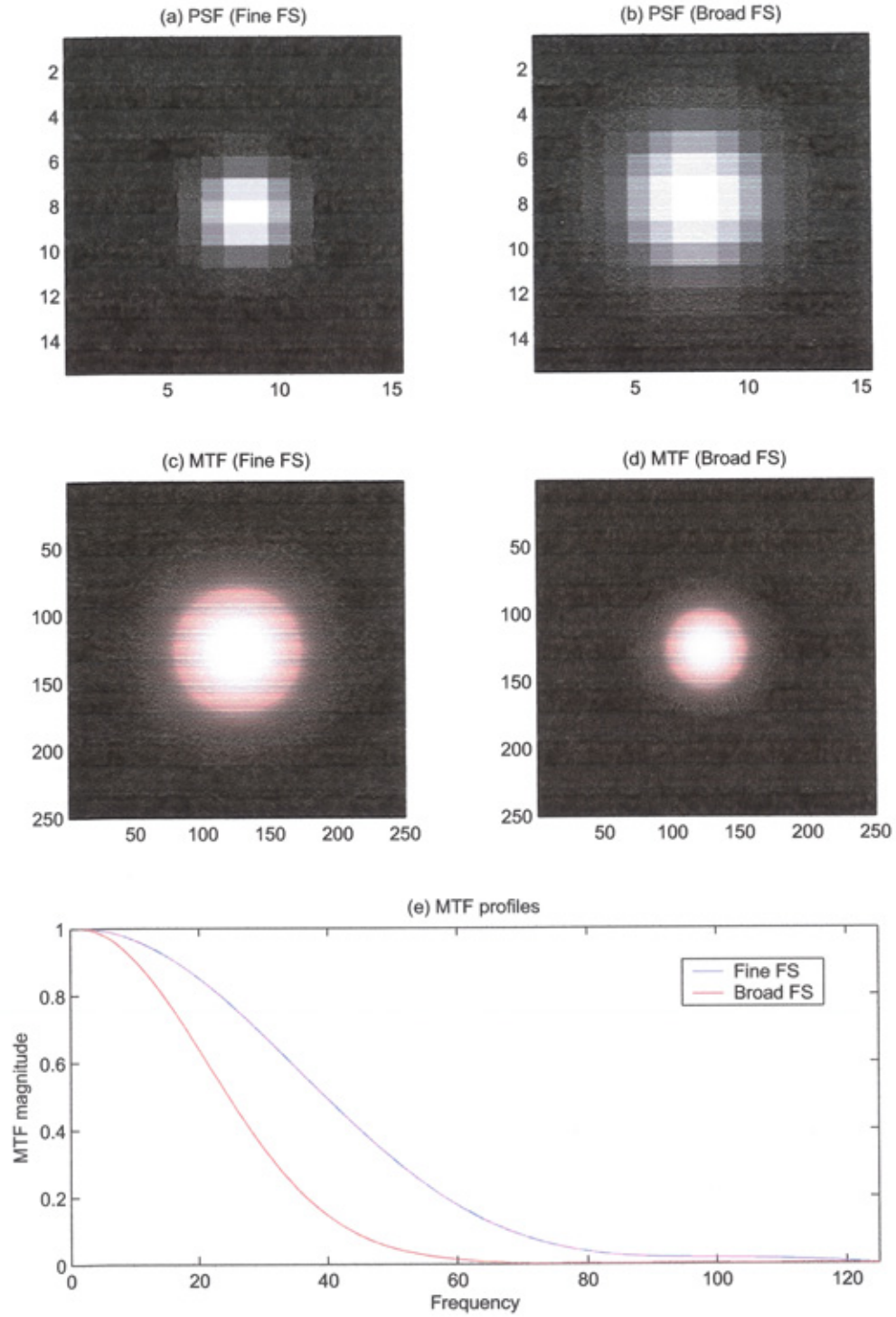

Figure 1.5: Images of fine and broad focal spots are shown in (a) and (b). The MTFs are shown below in (c) and (d) respectively. Profiles through each MTF are compared in (e). For increasing frequency (which is inversely proportional to spatial resolution) the MTF of a system with a broad focal spot is worse at retaining image power than a fine focal spot system.

values within that region. As the signal is approximately constant within the region any variations in pixel values is due to the noise.

Contrast resolution and windowing in CR

The idea of radiographic contrast was discussed earlier (see page 6 and section 1.3.1). Contrast resolution is the ability of a CR image receptor to discern contrast differences. The number of bits per pixel required for accurate representation of a monochrome image is often quoted as being between 6 and 7-bits (Castleman 1979; Pratt 1991; Gonzalez and Woods 1992). There is even some research (Cowlinshaw 1985) suggesting that as little as 4-bits (16 gray levels) per pixel is enough for most monochrome images on modern computer CRT screens. The luminance sensitivity of CR image receptors, which have a contrast resolution of 10, 12 or 14-bits (corresponding to 1024, 4096 or 16384 distinct shades of gray) far outstrips the performance of the human observer. Some manipulation of the acquired image is necessary in order to view the complete dynamic range of the image and this is usually achieved through windowing and/or contrast-stretching.

The very wide dynamic range of CR means that very low and very high signal strengths can be recorded in the same image. The CR system used to obtain many of the images in this work acquires images with 10-bit (1024) shades of grey. Modern computer displays are capable of showing 8-bit (256) shades of gray at a time. Therefore, some manipulation of the image is usually necessary to exploit the full benefit of the wide dynamic on an 8-bit display. We illustrate this with an example.

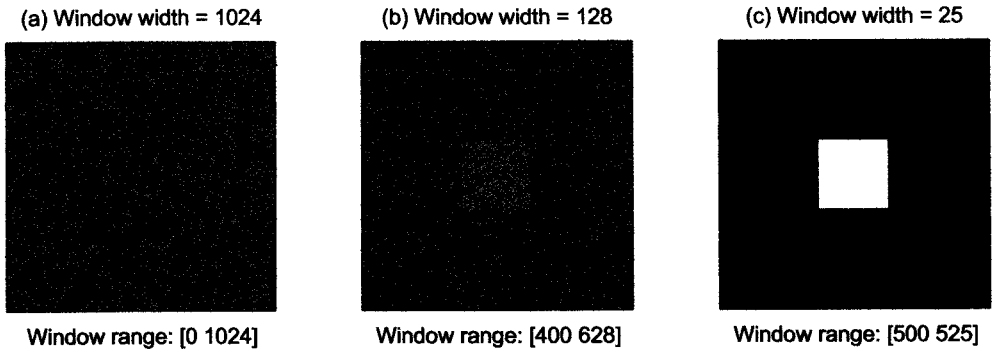

Figure 1.6: Demonstration of windowing to improve displayed contrast. The square feature has 5% more contrast than the background in terms of the original pixel values. Windowing can improve the visibility of this feature.

Figure 1.6 shows an image with potentially 1024 shades of gray. It depicts a square with pixel intensities of 525 against a uniform background with pixel intensities of 500. Contrast is defined in terms of intensity difference between the object and background: the contrast C between two regions with signal intensities I_A and I_B is given by:

$$C = |I_A - I_B| \tag{1.6}$$

In our example, the measured contrast difference between the object and background is 25; the object signal 5% higher than the background signal. Suppose we look at the image through 'windows' of varying widths. With a window width of 1024 the window width is equal to the potential dynamic range of the image. The 1024 gray levels are (linearly) mapped into 256 gray levels, shown in figure 1.6a. The object is difficult to distinguish from the background. If we decrease the window width to 128 gray levels we are free to position the window somewhere within the full 1024 grayscale range of the original image. Figure 1.6(b) shows the effect of positioning this window to view pixel values between 400 and 628 in the original image. Pixel values of 400 or less are shown as black, while pixel values of 628 or more are white. Values between 400 and 628 are mapped into the 8-bit range of the monitor display. Using a smaller window greatly improves the visibility of the square against the background, effectively increasing the displayed contrast. Figure 1.6(c) shows the effect of using a window width of 25 and positioning it so that the background is displayed at the bottom of the window, i.e. pixel values less equal to or less than the background are shown as black and the pixels values greater than or equal to the square are shown as white. The displayed image of the square has a greatly increased contrast.

Contrast, signal and noise

Contrast, as defined in equation (1.6), is a difficult property to quantify in a noisy image because it is not always clear where the contrast originates; it could be due to real differences in the attenuating properties of the tissue, or it could be due fluctuations in the noise. A human observer's ability to discern objects is proportional to contrast and inversely proportional to noise.

The contrast to noise ratio (CNR) is therefore an appropriate measure of the visibility of an image region. Substitution of equation 1.6 into the definition of CNR leads to:

$$\mathrm{CNR}_{AB} = \frac{\text{Contrast}}{\text{Noise}} = \frac{C}{\sigma} = \frac{|I_A - I_B|}{\sigma} = |\mathrm{SNR}_A - \mathrm{SNR}_B| \tag{1.7}$$

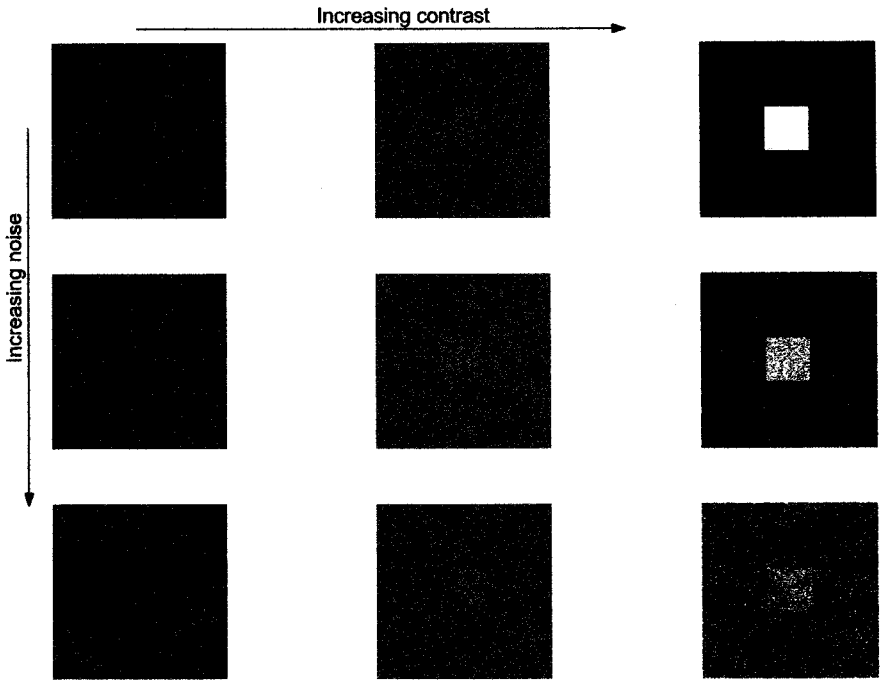

Figure 1.7: Demonstration of the effect of noise and contrast on feature visibility. The contrast of a digital image can be manipulated to aid detection of particular features (top row). However, this cannot overcome increasing levels of noise (bottom two rows).

The CNR between two regions is defined as the difference between the respective SNR of those regions.

Detective Quantum Efficiency

The ability of an image receptor to detect the radiation impinging on it is quantified by the Quantum Detection Efficiency (QDE). For example, the storage phosphor plates used in CR will not detect all the x-rays falling on it - some photons will be lost in transmission.

The ability of an image receptor to transform the detected photons into an image without introducing further noise is quantified by the Detective Quantum Efficiency (DQE). DQE therefore describes how much noise is added by the imaging system to the inherent quantum noise of the image. A perfect system would have a DQE of 100% of the QDE (no additional noise added to the quantum noise of the image) but in reality it may be much less than this.

Direct measurement of the DQE is not a straightforward task as it varies with the x-ray spectrum.

1.5 Some radiological applications

In previous sections it has been shown that the quality of medical images is susceptible to different forms of degradation such as geometric unsharpness and quantum noise. We now introduce two radiological applications to which image processing techniques might be applied in order to improve image quality.

1.5.1 Magnification mammography

X-ray mammography is a radiographic technique to examine the tissues of the breast. Mammography is a very technically demanding area of radiological imaging. The small differences in attenuation between breast tissues make adequate subject contrast difficult to obtain. The physical properties (atomic number, tissue density) of normal and malignant tissue are very similar. Low x-ray energies are employed to exploit these differences. However, these low energy x-rays are particularly susceptible to being absorbed by the breast and contributing nothing to the image. Mammographic imaging requires good spatial resolution to detect microcalcifications (discussed further below).

Mammography has been in widespread use since Egan adapted high-resolution industrial film to enable the production of mammograms with improved image detail in the 1960s. Direct exposure film was used to obtain those early mammograms and a high radiation dose to the breast tissues was necessary to obtain good quality images. Screen-film mammography was introduced in the 1970s and the improved technology allowed reduced radiation doses and higher contrast images to be produced. The evolution of modern mammography and related technology can be traced in the work of Jones (1982), Säbel and Aichinger (1996) and Bushberg et al. (2002).

Conventional screen-film contact mammography, in which the breast is placed close to the image receptor, is hampered by a low spatial resolution. An improvement to this technique is that of *magnification mammography* where the breast is positioned closer to the focal spot in order to geometrically magnify size on the image receptor. Funke et al. (1998) suggest that magnification mammography with CR may be superior to conventional screen-film mammography in microcalcification detection. An example of a mammogram is shown in figure 1.8.

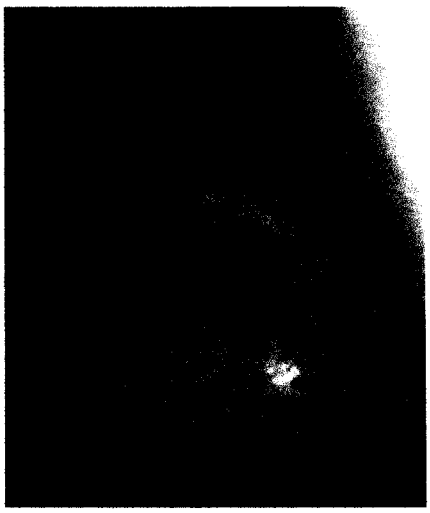

Figure 1.8: CR image of a breast obtained with contact mammography.

Breast cancer is the most common cancer in England with a rate of 7.3 cases per 1000 women in 2002/3 (Government Statistical Service 2004). The introduction of a national breast screening program in 1988 has meant that earlier detection has been possible in more cases and, combined with improved treatments, survival rates have risen. X-ray mammography has proved to an effective way of reducing breast cancer mortality rates in a number of other screening programs (Zhou and Gordon 1989; Hurley and Kaldor 1992). However, breast cancer remains the most common cause of cancer death in women.

Microcalcifications

Microcalcifications are small deposits formed by biological processes such as cell secretion or from necrotic cellular debris. The presence of microcalcifications in the breast can be an indicator of malignant processes at work in the breast some time before a carcinoma is seen (Hermann et al. 1988).

The sizes of microcalcifications typically fall into the range 0.1mm to 1mm. The composition of microcalcifications was investigated by Fandos-Morea et al. (1988) by analyzing the crystallographic patterns of microcalcifications. They found microcalcifications varied in composition between calcium carbonate, calcium oxalate and apatite (which is a phosphate mineral containing fluoride and phosphate anions).

The distribution of microcalcifications can give an indication of whether

a malignancy could form. Clusters of microcalcifications are more likely to arouse suspicion than scattered microcalcifications. The shapes of individual microcalcifications can also be an indicator; smooth, rounded microcalcifications are typically associated with benign processes whereas small, irregular shaped microcalcifications can be a sign of malignancy.

Mammographic imaging requirements

Obtaining good quality mammograms is a very technically demanding problem and specialised mammography units are used for the task. Some specific differences in mammographic imaging to the general principles of radiography already described are discussed below.

Low energy x-ray beam Radiography of the skull, chest or pelvis requires operating voltages of 60-110kVp. Mammography units are operated at lower voltages (25-35kVp) with 28kVp being usual. A low energy beam allows the production of an x-ray spectrum capable of generating images with good radiographic contrast.

Mo/Mo tube The most common tube design uses a molybdenum anode and filter, although rhodium filters are used in some circumstances. The lower energy x-ray photons used to acquire images with good radiographic contrast in mammography can be obtained using a molybdenum anode (rather than tungsten). Operating at a voltage of 24-30kV, a molybdenum filter strongly suppresses photons with energies in excess of 20keV which would otherwise reduce image contrast. The filter also absorbs low energy photons (<17.4keV[‡‡]) which would otherwise be absorbed by the skin and breast tissues and contribute nothing to the resulting image (Jones 1982).

Compression device Mammography units have devices to compress the breast during the exposure. Compression minimises the effects of scattered radiation (and thereby improving image contrast) by reducing the path length of the beam. Reduced scatter allows lower radiation doses. Compression also serves to reduce geometric and motion unsharpness. A balance must be found between enough compression and the comfort of the patient.

Automatic exposure control A sensor is placed behind the image receptor to trigger a shutdown of the x-ray source when the appropriate dose for the image has been reached.

[‡‡]This is K_α radiation energy of Molybdenum.

Geometric configuration In normal radiography the object is close to the image receptor, producing a 1:1 image. In mammography some degree of magnification may be required to aid detection of microcalcification clusters. In the conventional mammography set-up the distance between the focal spot and image receptor is usually approximately 60cm. Significantly shorter distances may result in a high dose of radiation to the skin. Magnification is achieved by placing the object approximately midway between the focal spot and receptor; this is illustrated in figure 1.2. A magnification of about 1.8 is used in practice.

Focal spot size In order for magnification to be useful the focal spot needs to be very small to minimise the blurring resulting from use of a finite sized focal spot.

Mammograms are generally obtained using screen-film image receptors rather than with CR or DR. However Cowen et al. (1997) show that the limiting resolution of CR does not translate into a dramatic difference in perception of microcalcifications at the limit of detectability. A recent account of CR in mammography is given by Seibert et al. (2004).

System optimisation of magnification mammography

The use of fine focal spots produces the clearest, sharpest images because geometric unsharpness caused by penumbral shadows is minimised. However, longer radiographic exposures are necessary for this method to be feasible due to the lower tube current limitations of the fine focal spot. Investigations into mammographic system optimization in terms of focal spot size, tube current, absorption efficiency and detector resolution have been carried out by Muntz et al. (1985). Conventional mammography configurations require focal spot sizes of less than $\approx$ 0.5mm while magnification mammography, which introduces greater geometric unsharpness, requires focal spots of less than $\approx$ 0.15mm (Medical Devices Agency 2001).

In this work we will explore an alternative approach which could prove useful in re-evaluating the system optimizations mentioned above. In the context of magnification mammography we will investigate the following ideas. It would be better to use a broad focal spot that allowed shorter exposures to x-ray radiation and less motion unsharpness caused by patient movement. As a consequence overheating the target of the x-ray machine would be also be less likely. However, the apparent advantages of a broad focal spot are more than offset by the loss of image quality caused by geometric unsharpness that would make magnification mammography self-defeating. There is a limit on how much magnification can be useful - even with a fine focal spot.

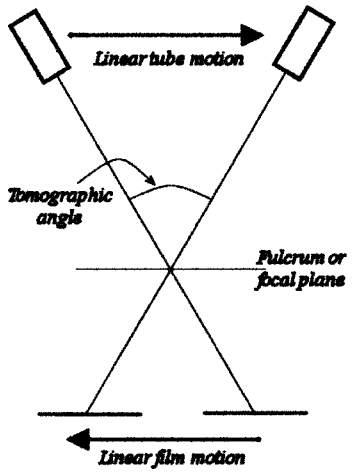

Figure 1.9: Simplified layout of a linear tomography unit.

Geometric magnifications higher than ≈ 2.0 produce unacceptable levels of blurring. The conventional set-up for magnification mammography is a fine focal spot with a geometric magnification around 1.8.

Image processing techniques would seem to offer some scope for improving images obtained using the conventional system set-up. In this research we will describe and test both conventional and unconventional set-ups. The aims of these tests are described below.

Conventional set-up To improve existing image quality or maintain image quality with reduced x-ray doses.

Unconventional set-ups To permit the use of higher magnifications and/or broad focal spots through reduction of the associated geometric unsharpness by image processing techniques. The possibility of dose reduction (increased quantum noise levels) will also be examined.

1.5.2 Linear tomography

Analogue tomography is a branch of radiography, originating in the 1930s, that uses motion of the x-ray focal spot and image receptor to elicit depth information from an object, which pre-dates modern digital tomography. In linear tomography, depth information is obtained by the opposing linear movements of the focal spot and image receptor (figure 1.9). In theory only a single plane within the object should remain in sharp focus while planes above

and below are blurred and, to some extent, removed from visual consideration by the human observer's eye-brain system.

Limited angle tomography and tomosynthesis

In limited angle tomography, rather than a obtain a single exposure while the focal spot and image receptor move in opposing directions, the x-ray tube is 'pulsed' to obtain a sequence of projection radiographs (rather than a single exposure as happens with conventional linear tomography). This process is called tomosynthesis. Any arbitrary plane in the object to be reconstructed retrospectively by 'shifting and adding' the individual radiographs. Tomosynthesis has advantages over conventional analogue tomography, where the focal plane is determined prior to the imaging process by the 'fulcrum' of the tomography machine.

Although the tomosynthesis image formation model is different from conventional tomography the problems of reduced visibility due to blurring from features outside the focal plane remain the same. The blurring can be improved mechanically by using motion more complicated than linear movement, such as circular or hypocycloidal. Image restoration methods have also been developed and applied to tomosynthesised images and we examine some of these approaches next. For a complete review of the development of digital tomosynthesis see Dobbins and Godfrey (2003).

Limitations of a linear tomography system

The images obtained with linear tomography correspond to 'slices' through the object and are not particularly thin. The thickness of a tomographic slice is inversely proportional to the tube swing angle and to some extent, the linear attenuation coefficient of the object (Hua et al. 1986). Linear tomography has largely been superseded by more complicated types of motion which produce less blurring and thinner 'slices'. Circular, figure-8 and hypocycloidal motion in particular may be better suited to a particular medical situation. However, there are still some situations in which the use of linear tomography is useful or essential. Examples include the evaluation of patients with nephrolithiasis (Goldwasser et al. 1989) and confirming diagnosis of allergic bronchopulmonary aspergillosis (Fisher et al. 1985).

The main limitation of linear tomography is that blurred features outside the focal plane may mask low contrast features within the focal plane of the object. Also, if a linear structure exists in the object and it is aligned with the linear motion, then it will not become blurred (except for a slight blurring at the ends) regardless of its depth in the object. The blurring

associated with linear tomography is complicated and involves distortions caused by differential magnification as well as motion blurring of those layers outside the focal plane. The shortcomings of linear tomography have meant applications in areas such as dentistry have moved towards tomography using, for example, spiral motion to reduce blurring. As an example we cite the work of Butterfield et al. (1997) which discusses the limitations of the clinical role of linear tomography in preoperative dental implant site assessment.

Nevertheless it is possible to model a linear tomography system and to seek 'de-blurred' solutions which approximate true 3-D data sets. This research will look at ways of removing the blurred features lying above and below the focal plane whilst retaining those features that are in focus and close to the focal plane. Presently this is something which can only be achieved at high cost using full scale CT-scanners.

CHAPTER 2

Image processing techniques

In this chapter a selection of popular image processing strategies will be examined and tested to determine the most robust method for dealing with images generated using the radiological applications described in Chapter 1. There are numerous methods which aim to improve the quality of image information content. It is possible to distinguish two different approaches to the problem; enhancement and restoration. It is also worth noting that the distinction between the two is somewhat blurred.

2.1 Enhancement *vs* Restoration

Image enhancement techniques are perhaps the most straightforward to implement. Enhancement techniques typically do not require a lengthy investigation into the causes of image degradation. The downside is that there is a considerable amount of subjective opinion as to when an image has been improved. Image enhancement does not retrieve any new information from an image but makes detail already in an image easier to see. This may be achieved in a number of ways, and involve manipulation of the image pixels or modifying a transformed version of the image (possibly in Fourier or wavelet space).

Image restoration techniques are generally more sophisticated than methods of image enhancement. They derive from an image formation model and seek to address and 'undo' the effects of image degradation directly. Degradation may take the form of missing data, blurring, noise or combinations of those effects.

In medical imaging there is scope to use image enhancement, image restoration or combinations of both.

2.2 Image enhancement methods

The appearance of some radiological images can be improved by the application of a suitable enhancement technique. The simplest methods of enhancement are functions which map pixel values in the input image to a new range in the output image. The purpose of this is to provide a better 'window' through which to view the image or a region of interest, or to increase the contrast between pixels; an example of this type of enhancement was seen in figure 1.6. Most commercially available software will provide these functions and are routinely used to view and manipulate digital radiographic images on-screen.

Another level of enhancement, beyond simple windowing of the image, is to manipulate the image information to extract or suppress particular properties of the image. Some useful (in a very general sense) enhancement techniques are:

Frequency filtering Filtering in the frequency domain to pass or suppress particular frequencies associated with features in the image.

High frequency emphasis filtering Modification of the frequency content of an image to increase the contribution of edges and fine structure.

Wavelet analysis Modification of the wavelet transform of an image to reduce noise at fine scales.

Each of these has corresponding disadvantages; methods which reduce noise also risk the loss of the fine structure which might be important for the radiologist to see. Conversely, increasing the contribution to the final image of fine structure or edges runs the risk of increasing the image noise and swamping the very features which the enhancement is trying to bring out. In recent years multiscale analysis (wavelets, ridgelets, beamlets, *etc*) has attempted to address these shortcomings by looking at the image components at different scales.

2.2.1 Frequency filtering

The Discrete Fourier Transform (DFT) is used to analyze or change the frequency content of an image (see Appendix A). An efficient implementation

of the DFT, called the Fast Fourier Transform (Cooley and Tukey 1965) is usually used in place of the DFT in image processing applications. All subsequent references to Fourier transforms in this work imply that the Fast Fourier Transform (FFT) was used for implementing the DFT.

The magnitude of frequency components* in a Fourier transformed image tends to decrease with increasing frequency; see figure 2.1(b) on page 34. Broad features and gradual intensity changes are characterised by the low frequencies of the Fourier transform and constitute most of the total power of the transform. Edges and sharp intensity transitions provide the strongest contributions to the high frequencies of the transform but the overall high frequency content of the transformed image is low. Uncorrelated noise is generally broadband and occurs at all frequencies but it tends to dominate the higher frequencies more easily.

Frequency filters modify the Fourier transform of an image. This is achieved by constructing a filter and multiplying it element-by-element with the Fourier transform of the input image. The filtered image is obtained by taking the inverse Fourier transform of the result. The simplest frequency filters block all low frequency content (high-pass filtering) or suppress the higher frequencies (low-pass filtering). More sophisticated filters aim to block or pass a narrow range of frequencies (band-pass or band-reject filters).

Ideal filters specify a cut-off frequency beyond which frequencies are either kept or suppressed. The sharp discontinuities introduced by ideal filters tend to contribute 'ringing' effects to the filtered image. This problem is overcome by constructing filters which have a smooth transition between the suppressed and passed frequencies. Gaussian and Butterworth filters do not have such discontinuities and are usually more effective than the ideal frequency filter. Examples of these filters are given in Gonzalez and Woods (2001).

A Matlab code to implement the frequency filtering discussed in this section is given in Appendix D.1.

Low-pass filtering

The application of a low-pass frequency filter to an image will remove the highest frequencies. Typically this will remove those frequencies dominated by noise and also those containing significant contributions from sharp intensity transitions (such as edges). Images which have been low-pass filtered are smoother, with the degree of smoothing depending on the cut-off frequency of the filter. The graininess associated with noise can often be removed but so too are other features with contributions from the cut-off frequency. Low-

*Sometimes referred to as *power* or equivalently as the MTF

pass filtering can be a useful first step for other applications such as line or edge detection but its appeal is somewhat limited in a radiological context, where reduction in noise is desirable but loss of image resolution is clearly not acceptable. The use of a low pass filter as a pre-processing step to image restoration is demonstrated in Chapter 3 (see figure 4.6 on page 97).

High-pass filtering

The high pass filter is especially useful in situations where the low frequency content is of little value. In linear tomography there is considerable blurring from features outside of the focal plane. This is manifested in the image as a substantial reduction in the high frequency content, in a direction corresponding to the direction of motion of the x-ray source and image receptor. A tomogram obtained from a linear tomography machine is shown in figure 2.1a. Below the tomogram is the corresponding Fourier transform, shown log-scaled in figure 2.1b. The blurring in the original tomogram is in a vertical direction and the effect is to remove fine structure and details in the vertical direction. In the Fourier transform of the image this is seen as a loss of power at high frequencies in the vertical direction. Image detail in the horizontal directions remain largely unaffected by vertical motion blur and so a greater range of frequencies is still to be seen across the width of the transform.

Filtering those frequencies whose contributions have been increased by the blurring - but only in the direction of the blurring - might be a useful way forward. Further investigation of high-pass filtering applied to linear tomography images are examined in Chapter 5.

2.2.2 Wavelets applied to radiological images

Wavelets are small waves of limited duration which satisfy certain mathematical conditions. In recent years wavelet analysis has been widely used in many areas of signal and image processing as an alternative to Fourier analysis.

Images can be decomposed into two or more scales using a wavelet transform. These scales contain information about the approximate image content (large features) and details (small features). An example of a wavelet decomposition for a noisy x-ray image is shown below.

In the following sections two applications of wavelets to radiological images will be described. The basic approach of both applications is the same and involves modification of the wavelet transform of the image in question:

1. Obtain the 2D wavelet transform of the image.

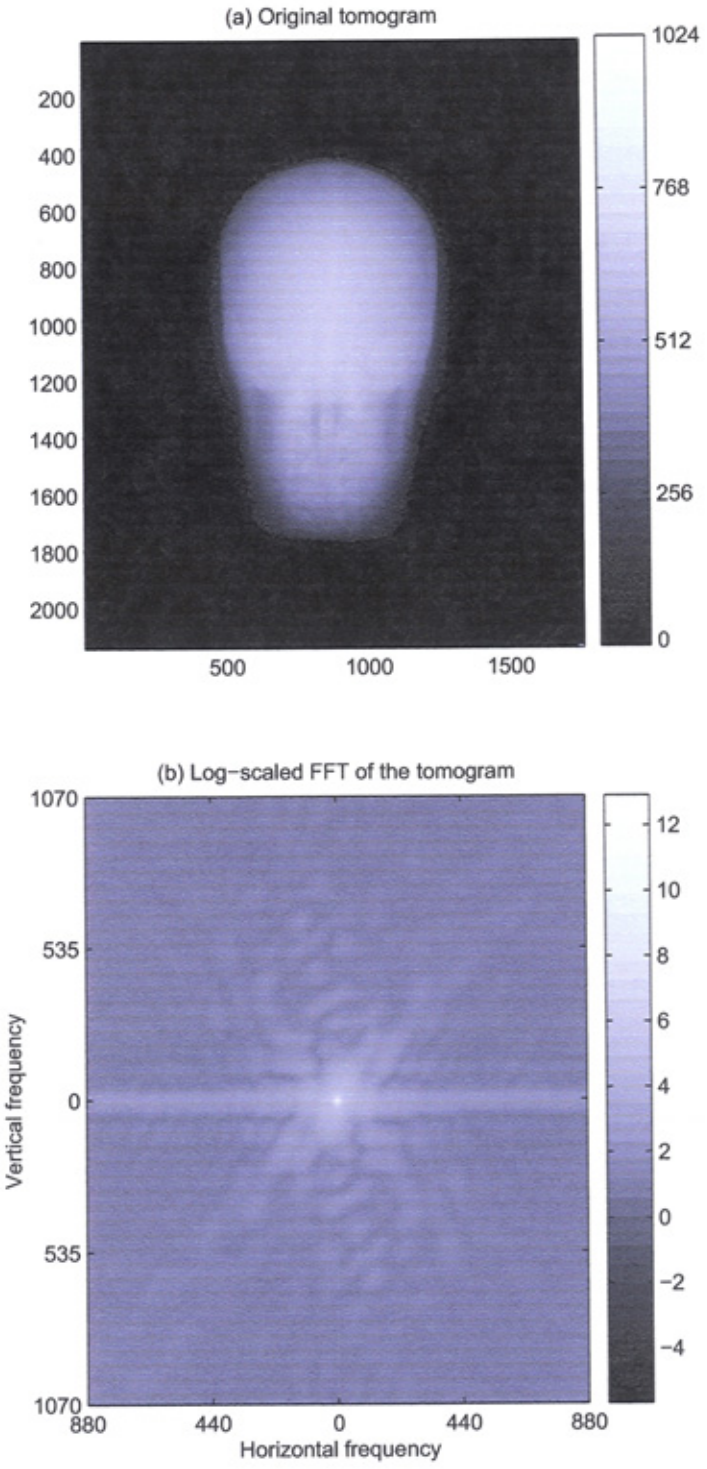

Figure 2.1: (a) is a tomogram through a plane in a skull phantom. The direction of motion in the image - and therefore blurring - was vertical. (b) shows the Fourier transform of (a). The blurring reduces the power of those frequencies corresponding to the vertical detail in the image; this is evident in the vertical direction of the transform.

2. Modify the transform.

3. Reconstruct the image from the inverse of the modified transform.

Denoising, enhancement and feature detection

The principal research into suppression of additive noise came from Donoho and Johnstone (Donoho 1992; Donoho and Johnstone 1992). Noise reduction is possible because the noise is typically seen as fine structure in the image and the wavelet transform allows decomposition of the image at different scales. Noise information is carried by wavelet coefficients at finer scales and by discarding them, a noise filtered image can be obtained by reconstructing from the modified transform. However, edge information is also represented by the fine scale wavelet coefficients so a threshold must be chosen in order to avoid disrupting desirable images features such as edges. Ideally, the threshold would separate noise related coefficients from edge related coefficients, allowing the latter to kept and the former to be discarded.

The type of thresholding procedure can be hard or soft. Hard thresholding is simplest and involves setting a value to which coefficients will be compared to and either kept or discarded depending on which side of it they fall. Hard thresholding can result in distortions within the reconstructed image because discontinuities are introduced in the wavelet transform. The alternative is soft thresholding, in which some of the coefficients above the threshold are scaled down towards zero. The distortions introduced by hard thresholding are thus avoided.

The reduction of noise in medical images can be problematical. The thresholding procedure described above is optimal under rather specific conditions. If applied without regard to the image formation process it may also remove 'noise' deemed to carry useful information in certain applications, such as speckle noise in an ultrasound images (Wagner et al. 1983). There are examples of denoising techniques applied to some medical images to be found in recent literature (Jansen and Bultheel 1997; Jansen et al. 1999; Pizurica et al. 2003; De Stefano et al. 2004; Ferrari and Winsor 2005).

We present our own examples of wavelet denoising in figure 2.2. The original image is from a phantom[†] containing fragments of bone imaged under low dose (2mAs), high noise conditions. In the first example, the image was decomposed with a Haar wavelet into two levels and soft thresholding was applied to the detail coefficients. The modified wavelet transform was reconstructed with the inverse wavelet transform to give the denoised image.

[†]This object, and further experiments using these images, will be described in Chapter 4.

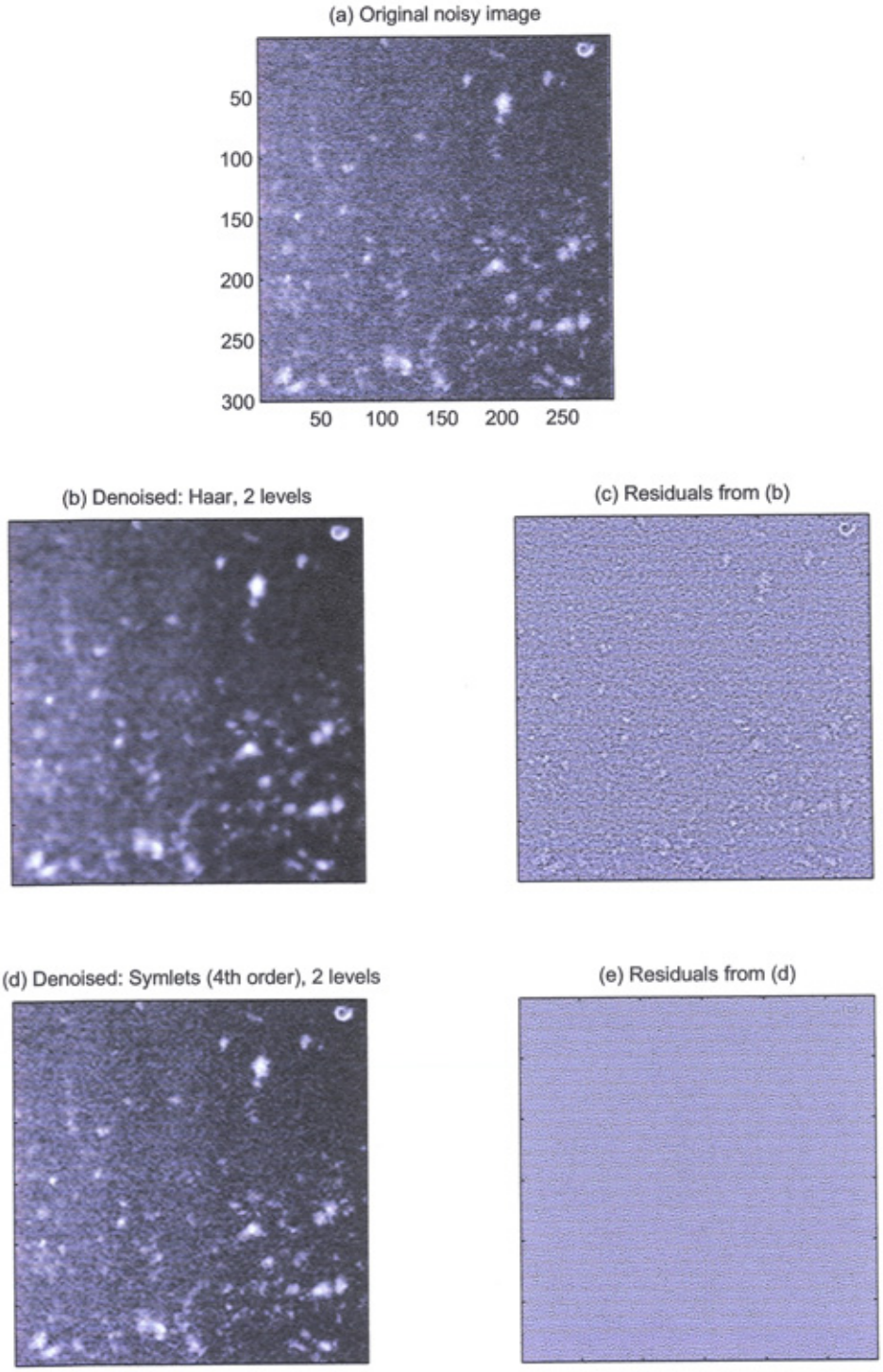

Figure 2.2: (a) is the original image of bone fragments imaged under low dose (high noise) conditions. Wavelet decompositions took place on 2 levels. (b) is a reconstruction denoised using a simple Haar wavelet. (d) is a reconstruction denoised using a 4th order Symlet wavelet. (c) and (e) show the information removed from the original in each case.

In the second example the same procedure was followed but using a 4th order Symlet wavelet. All processing was carried out using the Wavelet Toolbox in Matlab.

The residual images (c) and (e) show the information removed by the denoising process. Much of the information shown in (c) is noise although it is obvious that the edges of some real features were also removed. Almost all of the information in (e) is noise.

The use of wavelets to enhance the appearance of microcalcifications in mammograms has been discussed by Strickland (1996). The typical approach is to decompose the image using a wavelet transform and give more weight to the coefficients in the finer scales; in this respect the method is similar to the Fourier based high-frequency emphasis approach (see page 133). Performing an inverse wavelet transform results in an image where the microcalcifications have more contrast than in the original image.

Wavelet based detection of microcalcifications is described by Wang and Karayiannis (1998). The image is reconstructed from a wavelet transform which has had the lowest frequency subband completely suppressed. With an appropriate choice of wavelet and when used in conjunction with a non-linear thresholding method, enhanced detection of groups of microcalcifications was possible in a mammogram.

More recently Sakellaropoulos et al. (2003) described a wavelet-based method of minimizing image noise while optimizing image contrast. This was achieved by modifying the multiscale gradient magnitude values obtained via the dyadic wavelet transform. Noise reduction was implemented with a spatially varying denoising technique similar to the method of Donoho discussed earlier. Contrast enhancement was attained by applying a local linear mapping operator on the denoised wavelet magnitude values.

2.3 Image restoration methods

Restoration methods are derived from an the image formation model of an imaging system. For a system where the degradation is completely characterised by the system's response to an impulse, i.e. the point-spread function $h(i, \alpha, j, \beta)$, the relationship between the input image $x(\alpha, \beta)$ and output image $d(i, j)$:

$$d(i,j) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x(\alpha,\beta) h(i,\alpha,j,\beta)\, d\alpha d\beta \tag{2.1}$$

Equation 2.1 is an example of a Fredholm integral equation of the first kind. This equation is at the heart of all subsequent image restoration meth-

ods. and further investigation of the properties of this integral equation will be discussed in chapter 3. This equation is at the heart of all subsequent image restoration methods. For the medical imaging problems considered in this work we will assume a linear, spatially invariant degradation function, *i.e.* $h(i,\alpha,j,\beta) = h(i-\alpha, j-\beta)$. In this case the above equation reduces to:

$$d(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(\alpha,\beta)h(i-\alpha, j-\beta)\, d\alpha d\beta \tag{2.2}$$

In the presence of random system noise which is not dependent on position in the image we can add a noise term to get:

$$d(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x(\alpha,\beta)h(i-\alpha, j-\beta)\, d\alpha d\beta + n(i,j) \tag{2.3}$$

This image formation model assumes that a 'true image' has been blurred by a point-spread function (PSF) and then further degraded by additive noise. This is a simplification of the processes which lead to the production of radiographic images. For example, it is implied in equation 2.3 that the PSF is spatially invariant and that the noise is independent of the image. In reality, neither assumption is entirely satisfactory. However, the formulation of many restoration methods is simplified by this approach and the computational expense is greatly reduced because the resulting matrix-vector equations can be solved through the use of fast Fourier transforms.

The images have thus far been treated as continuous functions but in reality we are dealing with digital images with discrete pixels. The discrete form of equation (2.3) is:

$$d_{ij} = \sum_{\alpha,\beta} x_{\alpha\beta} h_{i-\alpha,j-\beta} + n_{ij} \equiv x * h_{ij} + n_{ij} \tag{2.4}$$

where '*' denotes the mathematical operation of convolution; the first term on the right hand side is called the forward map. In medical radiography, we are given the output image data d_{ij}, some knowledge of the noise and information describing the degradation (point-spread) function h. We would like to obtain an image $\hat{x}_{ij}$, which is an approximation to the 'true' image, x_{ij}, to within the limits set by the noise and data. A diagram of this model for a radiological imaging system is shown in figure 2.3.

Image restoration techniques can be less subjective than enhancement techniques as the images produced will satisfy some *a priori* definition of restoration. In the sections that follow, a selection of deconvolution techniques will tested to examine their effectiveness.

1. Algebraic deconvolution

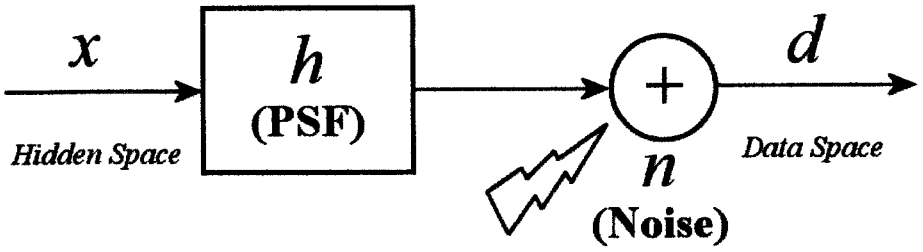

Figure 2.3: Image formation model.

2. Inverse and pseudoinverse filtering
3. Wiener filtering
4. Regularised deconvolution
5. Other methods

This is not a comprehensive list of all available deconvolution methods; the purpose is examine the shortcomings of methods which have previously been applied to problems of deblurring medical images. The types of images obtained in medical radiography are large, typically with dimensions of several millions of pixels. Medical images are noisy (for the reasons described in Chapter 1). The examinations of algebraic deconvolution (1) provides clues to whether or not it is possible to extract solutions to these imaging problems, and whether the solution obtained is unique or from a space of many, or infinitely many possible solutions. Restoration techniques are most efficiently implemented using Fourier transforms; (2) and (3) are techniques which implement these restorations quickly. (4) attempts to select a solution from the myriad of possible solutions by minimising some criterion of smoothness with the constraint that the solution is consistent with the observed image. There are many possible criteria for image smoothness and it is the aim of this chapter to show that one property in particular - the entropy - is the best choice for deconvolution of medical images. Finally, in (5) we mention some notable absences; methods of deconvolution used in other image deblurring problems and why their use was not considered appropriate for the problems discussed in this work.

2.3.1 Algebraic deconvolution

Algebraic deconvolution is the analysis and solution of the image restoration problem with systems of equations represented by matrices and vectors. The

image formation model (equation 2.4) can be represented algebraically by:

$$\mathbf{d} = A\mathbf{x} + \mathbf{n} \tag{2.5}$$

where A is a matrix representation of the point-spread function h_{ij} from equation 2.4. The vectors $\mathbf{x}$, $\mathbf{d}$ and $\mathbf{n}$ are formed by stacking columns of the corresponding images x_{ij}, d_{ij} and n_{ij}. The goal of restoration is to find a solution $\hat{\mathbf{x}}$ such that:

$$A\hat{\mathbf{x}} = \mathbf{d} \tag{2.6}$$

Assuming for the moment that A is square and nonsingular, by substituting equation (2.6) into equation (2.5) and applying the inverse of A to both sides we obtain:

$$\begin{aligned} A\hat{\mathbf{x}} &= A\mathbf{x} + \mathbf{n} \\ A^{-1}A\hat{\mathbf{x}} &= A^{-1}A\mathbf{x} + A^{-1}\mathbf{n} \\ \hat{\mathbf{x}} &= \mathbf{x} + A^{-1}\mathbf{n} \end{aligned} \tag{2.7}$$

So the restoration $\hat{\mathbf{x}}$ can be considered as being the sum of two components. The first term is the true image distribution and the second is a noise amplification term. The magnitude of the noise amplification term would seem to depend on the properties of the matrix A. The inverse of the matrix is given by:

$$A^{-1} = \frac{\text{adj}(A)}{\det(A)} \tag{2.8}$$

where 'adj' and 'det' are the adjoint and determinant of the matrix A respectively. If the matrix A is ill conditioned then $\det(A)$ will be very small and the inverse A^{-1} will have some very large values. In these situations the noise amplification term will dominate equation (2.7). If A is singular the situation is worse; the determinant is equal to zero and A^{-1} does not exist at all. In this case equation 2.6 has many solutions (if $\mathbf{d}$ is in the column space of A) or no solutions (if $\mathbf{d}$ is not). We will show in chapter 3 that the solutions to image restoration problems are at best ill-conditioned and at worst, they are singular.

Matrix representation of the PSF

Algebraically, the convolution of the true image and PSF ($h_{ij} * x_{ij}$) can be implemented as an equivalent matrix-vector multiplication ($A\mathbf{x}$), where A is

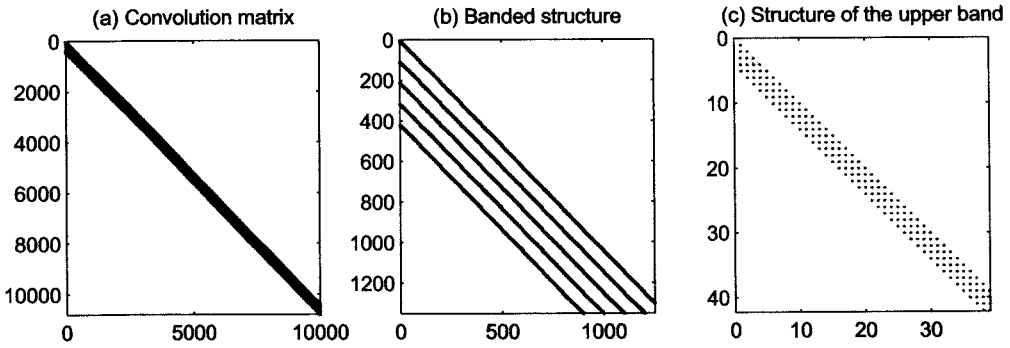

Figure 2.4: The sparsity and structure of the matrix representation of a 5×5 averaging PSF, applicable to an image with dimensions of 100×100. (a) shows the complete matrix. (b) shows the banding structure towards the upper-left. (c) Shows the structure of the uppermost band in (b). In all three cases non-zero matrix elements are denoted by a blue dot.

a matrix representation of h_{ij}. The procedure for constructing A from h_{ij} is described in detail by Gonzalez and Woods (1992). We assumed in our earlier analysis that the matrix A was square and non-singular. In practice A is a non-square Block-Circulant matrix whose construction procedure incorporates *image padding*, i.e. references to pixels beyond the image border. Figure 2.4 shows a plot of two convolution matrices showing the sparse structure of these objects. Matrix elements which are non-zero are shown with a blue dot - everywhere else is zero. In our example, of the 108 million elements in A only 250,000 (0.0023% of the total) are non-zero.

Non-square matrices correspond to over-determined or under-determined linear systems. In these cases the matrix A does not have conventional inverse. Approximate solutions to equation 2.5 may be sought through other methods, such as the singular value decomposition and generalised inverse.

Singular value decomposition

The singular value decomposition (SVD) of a general $m \times n$ matrix A is:

$$A = V\Sigma U^T \tag{2.9}$$

where the matrix V is $m \times m$, Σ is $m \times n$ and U is $n \times n$. This decomposition is obtained by calculating the eigenvalues λ_i, and the associated eigenvectors $\mathbf{u}_i$, of $A^T A$. This matrix is square and symmetric, so that its eigenvalues are real. It can also be shown that they are also non-negative, i.e. $\lambda_1 \ldots \lambda_n \geq 0$;

k of these will be strictly positive, and can be ordered so that $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_k \geq 0$, where $k \leq n$. The first k columns of the matrix V are the corresponding m-dimensional vectors $\mathbf{v}_i$ defined by

$$\mathbf{v}_i = \frac{1}{\sigma_i} A\mathbf{u}_i \tag{2.10}$$

where $\sigma_i = \sqrt{\lambda_i}$, $i = 1 \ldots k$. The remaining columns of V are any $\mathbf{v}_{k+1} \ldots \mathbf{v}_k$ such that the full set $\mathbf{v}_1 \ldots \mathbf{v}_m$ is orthonormal. The columns of U are the (normalised) eigenvectors $\mathbf{u}_i$. Also, U^T and V are orthogonal matrices.

The matrix Σ has the special form

$$\Sigma = \begin{pmatrix} D & 0 \\ 0 & 0 \end{pmatrix} \tag{2.11}$$

where:

$$D = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_i \end{pmatrix} \tag{2.12}$$

The numbers $\sigma_1, \ldots, \sigma_i$ are all real, positive and greater than zero. Collectively these are referred to as singular values of the matrix A. Also, the singular values form a numerically decreasing sequence $\sigma_1 > \sigma_2 > \ldots > \sigma_i > 0$.

The *condition number* of a matrix A can be specified in terms of the maximum and minimum singular values:

$$\text{cond}(A) = \frac{\sigma_{max}}{\sigma_{min}} \tag{2.13}$$

The matrix A is said to be *well-conditioned* if the condition number is small (close to 1) and *ill-conditioned* if it is very large.

Generalised inverse

By SVD we can write $A = V\Sigma U^T$ for any $m \times n$ matrix. If the matrix is square ($n \times n$) and Σ is invertible (i.e. all of the diagonal elements - the singular values - are non-zero) then the inverse is simply

$$A^{-1} = (V\Sigma U^T)^{-1} = U\Sigma^{-1}V^T \tag{2.14}$$

since U and V are orthogonal. When A is not square then Σ is also not invertible and A^{-1} does not exist. In this situation the best that can be done is to construct a *generalised inverse*

$$A^+ = U\Sigma^+V^T \tag{2.15}$$

where Σ^+ is the $n \times m$ matrix

$$\Sigma^+ = \begin{pmatrix} D^{-1} & 0 \\ 0 & 0 \end{pmatrix} \tag{2.16}$$

Equation 2.15 is called the Moore-Penrose inverse[‡] (Moore 1920; Penrose 1955) and is the most commonly used of several generalised inverses. It is this matrix inverse to which we refer in subsequent discussions. The Moore-Penrose inverse is a unique matrix inverse which satisfies certain mathematical conditions: for example, see Hill (1996).

The solution $\hat{\mathbf{x}}$ which approximates the true image distribution $\mathbf{x}$ in equation (2.6) is obtained from:

$$\hat{\mathbf{x}} = A^+\mathbf{d} \tag{2.17}$$

It is that solution which simultaneously minimizes the error $\|A\hat{\mathbf{x}}-\mathbf{d}\|$ and the norm $\|\hat{\mathbf{x}}\|$. In an underdetermined system the minimum error will be zero. We shall illustrate the SVD and generalised inverse with a simple example. Consider the 2×3 matrix

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$$

giving

$$A^TA = \begin{pmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{pmatrix}$$

which is obviously singular. The eigenvalues and corresponding normalised eigenvectors are

$$\begin{aligned} \lambda_1 = 6 \quad & \mathbf{u}_1 = \tfrac{1}{\sqrt{3}} \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T \\ \lambda_2 = 0 \quad & \mathbf{u}_2 = \tfrac{1}{\sqrt{6}} \begin{pmatrix} 2 & -1 & -1 \end{pmatrix}^T \\ \lambda_3 = 0 \quad & \mathbf{u}_3 = \tfrac{1}{\sqrt{6}} \begin{pmatrix} 0 & 1 & -1 \end{pmatrix}^T \end{aligned}$$

giving

[‡]The Moore-Penrose inverse is sometimes referred to as the pseudoinverse. However, to avoid confusion with a Fourier based restoration technique introduced later in this chapter, we will refrain from associating this term with the Moore-Penrose inverse.

$$U = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix}$$

Thus $k = 1$, giving

$$\mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

as the only orthonormal choice; thus

$$V = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

Thus

$$\Sigma = \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\Sigma^+ = \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$$

The generalised inverse is therefore

$$\begin{aligned} A^+ &= U\Sigma^+V^T \\ &= \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \\ &= \frac{1}{6} \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \end{aligned}$$

The system

$$A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{b} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

is underdetermined, as the vector $\mathbf{b}$ is in the column space of A; according to equation 2.17 the required solution is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2/3 \\ 2/3 \\ 2/3 \end{pmatrix}$$

which is exact. It is obvious that $(\frac{2}{3}, \frac{2}{3}, \frac{2}{3})$ is the point in the plane $x+y+z = 2$ which is closest to the origin. If we choose $\mathbf{b} = (1,0)^T$ the system is overdetermined, and the required solution is

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/6 \\ 1/6 \\ 1/6 \end{pmatrix}$$

which is not exact. The point $(\frac{1}{6}, \frac{1}{6}, \frac{1}{6})$ is the point in the plane $x+y+z = \frac{1}{2}$ closest to the origin.

Use of the Moore-Penrose inverse is not restricted to problems in which A is singular. Ill-conditioned problems (when the condition number of A is very large) can be stabilised by truncating the matrix D (equation 2.12) at row k beyond which the singular values $(\sigma_{k+1} \ldots)$ on the diagonal are significantly smaller than σ_1.

Practical issues with algebraic deconvolution methods

It is worth noting that if the original image d has k^2 pixels then algebraic deconvolution requires the manipulation of a PSF matrix A of dimensions $k^2 \times k^2$. For images containing more than a few thousand pixels the computational effort involved in calculating the SVD, say, of the PSF matrix is enormous. Although processing power has improved greatly in recent years, algebraic deconvolution methods are still computationally very expensive. Algebraic deconvolution methods dealing with large images try to exploit the sparsity and structure of A. In some practical applications A is never formed or stored in its entirety, but rather approximated by a Kronecker product of two smaller matrices. For a more detailed description of these methods see Nagy (1996) and Kamm and Nagy (1998, 2001). A general review of these and other algebraic deconvolution techniques is described by Hansen (2002).

Matrix-vector multiplications in which the matrix is Block-Circulant are more quickly implemented in the frequency domain using the Discrete Fourier Transform; an explanation is given in Gonzalez and Woods (1992). Fourier based methods form the basis of the restoration techniques discussed in the remainder of this chapter.

The practical difficulties associated with algebraic deconvolution give these methods limited appeal to large-scale computational problems. However, a literature search shows interesting applications to medical deconvolution: Yuasa et al. (1997) describe an attenuation correction method for detection of small amounts of nonradioactive contrast material used in fluorescent x-ray CT, and Koh et al. (2004) used SVD in the assessment of tissue perfusion by dynamic contrast-enhanced imaging.

2.3.2 Test image: simulated blur and noise

In the analysis which follows an clinical test object (a 'phantom') was used to evaluate the performance of restoration methods. The CDMAM phantom consists of an aluminium base with gold disks of different sizes and thicknesses attached to a plexi-glass cover. The image in figure 2.5 was obtained with the magnification 1.8 fine focus geometric set-up and high dose (low noise) conditions. For the purposes of testing the methods that follow, figure 2.5a is considered to be the 'true image'. The true image in this case is a region with dimensions (in pixels) of 497 × 484 which was cropped from a much larger full-size image with dimensions of 2370 × 1770. The 'true image' also contains a few artefacts[§] in the form of bright pixels.

The image under study in these tests is a 10-bit DICOM compatible image: each image potentially contains 1024 shades of grey. Image processing was carried out with Matlab, which treats the images as matrices. Prior to processing the image matrices were converted from integers to double precision numbers in the range [0 1] by dividing through each element by 1024.

Blurring and noise were subsequently added using Matlab's Image Processing Toolbox. The type of blurring associated with mammography is simulated in figure 2.5b. The scaled PSF of a real focal spot was used to generate the blurred image. This image was then further degraded by medium and high levels of noise in figures 2.5c and 2.5d respectively.

2.3.3 Inverse and pseudoinverse filtering

In the frequency domain, for an imaging system free from noise the observed image data is related to the true image by the equation:

$$D_{uv} = X_{uv} H_{uv} \tag{2.18}$$

[§]Described in more detail in Chapter 4, page 98

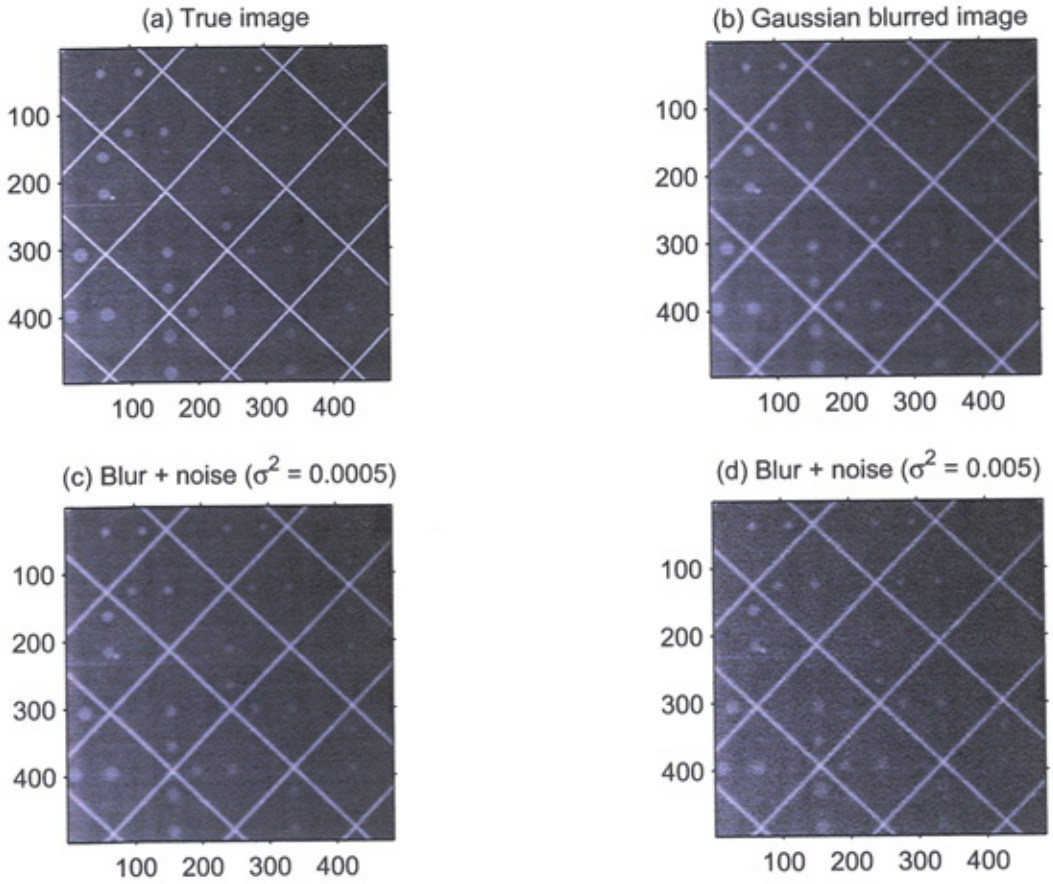

Figure 2.5: A section of the CDMAM phantom. (a) is considered to be the 'true image' in subsequent analysis, (b) results when the true image is blurred with a Gaussian PSF 9 pixels wide. If we add some Gaussian noise we obtain (c) and (d) with medium and high levels respectively.

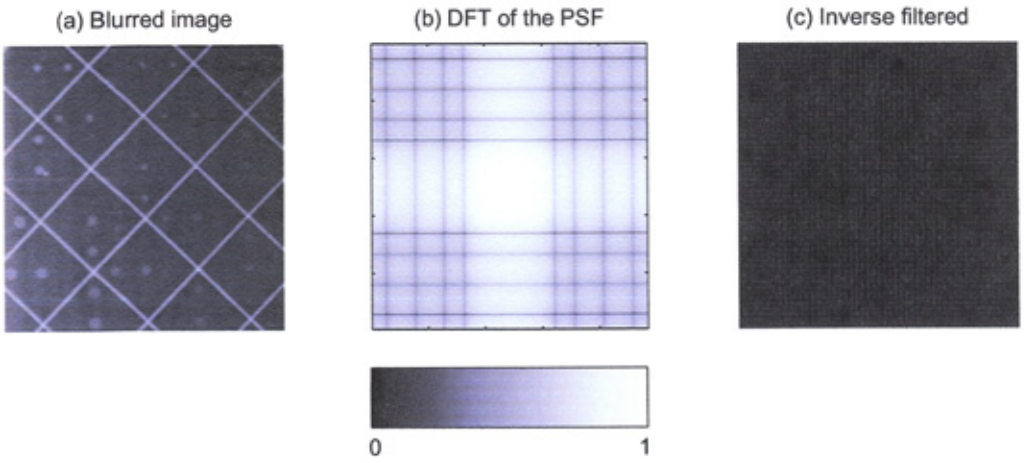

Figure 2.6: The inverse filter fails because there are many small and zero values (seen criss-crossing the image) in Fourier transform of the PSF. The Fourier transform shown in (b) has been log-scaled for greater clarity.

where the capitalised quantities here are obtained by Fourier transforming the lowercase quantities from equation 2.4. The equivalence of equations 2.4 and 2.18 form the *convolution theorem*. To obtain an approximation of the transform of the 'true' image we might simply rearrange this equation to get:

$$\hat{X}_{uv} = \frac{D_{uv}}{H_{uv}} \tag{2.19}$$

Application of the inverse Fourier Transform should give us the 'true' image in the spatial domain. This approach is known as inverse filtering. In the absence of noise then any zeros in H_{uv} should coincide with zeros in D_{uv}, according to equation 2.18. The ratio 0/0 is of course undefined mathematically. It is obvious from equation (2.19) that the inverse of any near-zero values of H_{uv} will tend to become very large indeed. Even when the PSF is well behaved this may still be the case.

The results of inverse filtering are shown in figure 2.6 and it is seen that the inverse filter completely fails to recover the image. The presence of very small and/or zero values in the frequency domain of the PSF is to blame for the failure of the method. This computational difficulty is overcome by thresholding and redefining the inverse of H_{uv} (Jain 1988):

$$H^I_{uv} = \frac{1}{H_{uv}} = \begin{cases} \frac{1}{H_{uv}} & H_{uv} \neq 0 \\ 0 & H_{uv} = 0 \end{cases} \tag{2.20}$$

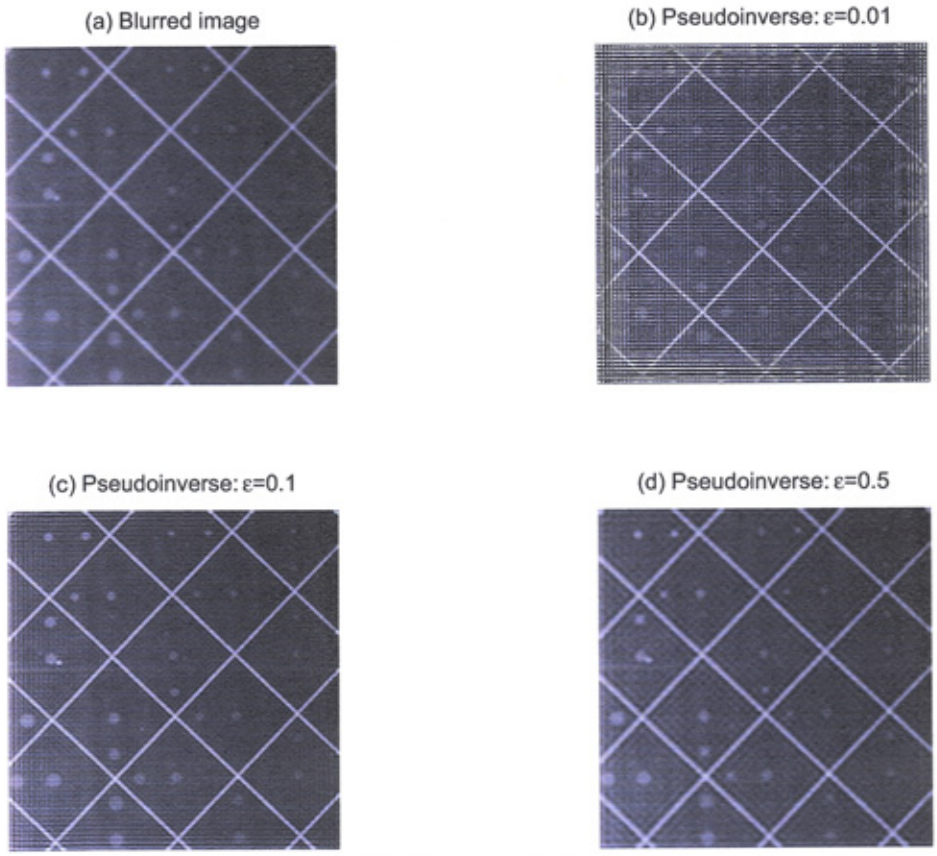

Figure 2.7: Pseudoinverse filtering can perform well but exhibits ringing effects as the restorations become sharper. If ε is increased too much then blurring begins to take over once more as the image is oversmoothed.

In practice H^I_{uv} is set to zero whenever $|H^I_{uv}|$ is less than a suitably chosen positive quantity ε.

$$H^I_{uv} = \frac{1}{H_{uv}} = \begin{cases} \frac{1}{H_{uv}} & |H_{uv}| > \varepsilon \\ 0 & |H_{uv}| \leq \varepsilon \end{cases} \tag{2.21}$$

As $\varepsilon \longrightarrow 0$ so equation (2.21) approximates more closely the true inverse filter of equation (2.20). This thresholding version of the inverse filter is also known a pseudoinverse[¶] filter. The true image figure 2.5a was blurred with a PSF with a Gaussian shape and a width 9 pixels. The results of inverse and pseudoinverse filtering are shown in figure 2.7. The Matlab code to implement these filters is given in Appendix D.3. The pseudoinverse filter produces acceptable restorations in these images although ringing artefacts begin to dominate as the restoration produces sharper images. If ε is too large the restored image begins to become blurred once more (as in figure 2.7d).

The image considered in this section contained no noise component and further tests showed that the pseudoinverse filter is not very robust when applied to images containing even small amounts of noise.

2.3.4 Wiener filtering

The degradation model for a blurred, noisy image has the following form in the frequency domain.

$$D_{uv} = X_{uv}H_{uv} + N_{uv} \tag{2.22}$$

Substituting this expression into equation (2.19) yields an approximation to the true image:

$$X_{uv} = \hat{X}_{uv} - \frac{N_{uv}}{H_{uv}} \tag{2.23}$$

It is obvious that any zero or near-zero values of H_{uv} will greatly amplify contribution from the noise term N_{uv}. The problem is that the noise spectrum N_{uv} is not usually known with sufficient accuracy to enable X_{uv} to be recovered, often compounded by the fact that H_{uv} has zero or very small values, at which points the ratio $\frac{N_{uv}}{H_{uv}}$ can easily dominate equation (2.23). This is related to the fact that the forward map *true image* $\longrightarrow$ *degraded image* is often singular or almost singular (many to one).

[¶]Again, we point out that this is unrelated to the Moore-Penrose matrix inverse discussed earlier, which is sometimes called a pseudoinverse.

In practical situations a restoration method that can deal with the problem of noise is required. The Wiener filter (Wiener 1949; Helstrom 1967) attempts to address some of the failings of direct Fourier inversion techniques when applied to noisy images. The Wiener filter is based on a statistical criterion which minimises the mean square error between the true image distribution x and the estimate of it $\hat{x}$, i.e.

$$e^2 = E\{(x - \hat{x})^2\} \tag{2.24}$$

The Wiener filter is also called the least square error filter and as such is optimal in an average sense. The Wiener filter approach gives an estimate $\hat{X}_{uv}$ of the true image:

$$\hat{X}_{uv} = \left[\frac{1}{H_{uv}} \frac{|H_{uv}|^2}{|H_{uv}|^2 + K_{uv}}\right] D_{uv} \tag{2.25}$$

where K_{uv} the inverse of the signal-to-noise ratio (SNR), often approximated by a constant. It can be thought of as an adjustable parameter chosen by the user to balance sharpness against noise in the restored image. This is a fairly subjective choice but it can work well in images where the SNR is not too low. The Wiener filter tries to remove those frequencies which would otherwise be dominated by noise in the Fourier transform of an image. To operate effectively the Wiener filter requires the PSF of the forward map to be known. Several assumptions about the image, noise and forward map are also made:

- The forward map is linear. In other words the pixel intensities in the observed image are linearly related to those in the 'true' image.
- The blurring is spatially invariant. This ensures that the PSF matrix (on which the underlying theory is built) remains block-Toeplitz in form. There is a small PSF variation across our radiographic images as the focal spot is 'viewed' from different perspectives, which we ignore.
- The image is uncorrelated. That is to say, pixel correlations are negligible beyond a certain small distance and that the image itself can be modeled as a wide-sense stationary stochastic process (*i.e.* one in which the underlying autocorrelation functions are spatially invariant).
- The noise is signal independent with zero mean. Again, this ensures that the assumptions about the autocorrelation matrices, from which the Wiener filter is derived, remain valid.

The Wiener filter derives from a linear model: application to a non-linear image formation process would be inappropriate. These assumptions lead to autocorrelation matrix approximations in Toeplitz and Block Toeplitz form. Such matrices are readily approximated by Block Circulant matrices and multiplications are more efficiently implemented using the Discrete Fourier Transform.
A Matlab code to implement the Wiener filter is listed in Appendix D.4. With reference to figure 2.8, the images on the left hand side are blurred and with increasing levels of Gaussian noise added. The results of Wiener filtering each image are shown on the right-hand side.

The Wiener filter performs reasonably well in images where the signal-to-noise ratio is high (i.e. the very low levels of noise). The restored image is very sharp but ringing artefacts remain throughout the image. As the levels of noise increase, the effectiveness of the Wiener filter decreases. Suppressing the noise by choosing larger values of K reduces the effectiveness of deblurring.

Wiener filtering and radiological images

The Wiener filter has been tested on simulated radiological images (Madsen 1990) and to 'real' radiological images (Dougherty and Kawaf 2001). In the latter case, the authors examined the restorative effectiveness of a commercially available implementation of the Wiener filter. The geometric set-up in their experiment was such that a magnification of 4.0 was obtained with a broad focal spot (0.6mm in this case). The image receptor was film (with the intensifying screen absent to reduce the associated blurring). The Wiener filtered result showed improved spatial resolution over the original image with some evidence of noise-smoothing. However, the restored images also exhibited 'ringing' artefacts close to high luminance edges.

SPECT[‖] is a tomographic technique using gamma ray emitting radiopharmaceuticals. Wiener filtering of 3D SPECT images to reduce collimator blur was described by Hawkins and Leichner (1994). The filter was applied between the pre-multiply and backprojection phases of reconstruction. The Wiener filter was therefore an integral part of the process rather than a pre or post-processing technique. The technique outperformed both post-processing with a Wiener filter and the quantitative CHT SPECT protocol (which utilises a 2D Wiener pre-filtering approach) in terms of noise amplification.

Our experiences with the artificially degraded images seen in this chapter,

[‖]SPECT: Single Photon Emission Computed Tomography

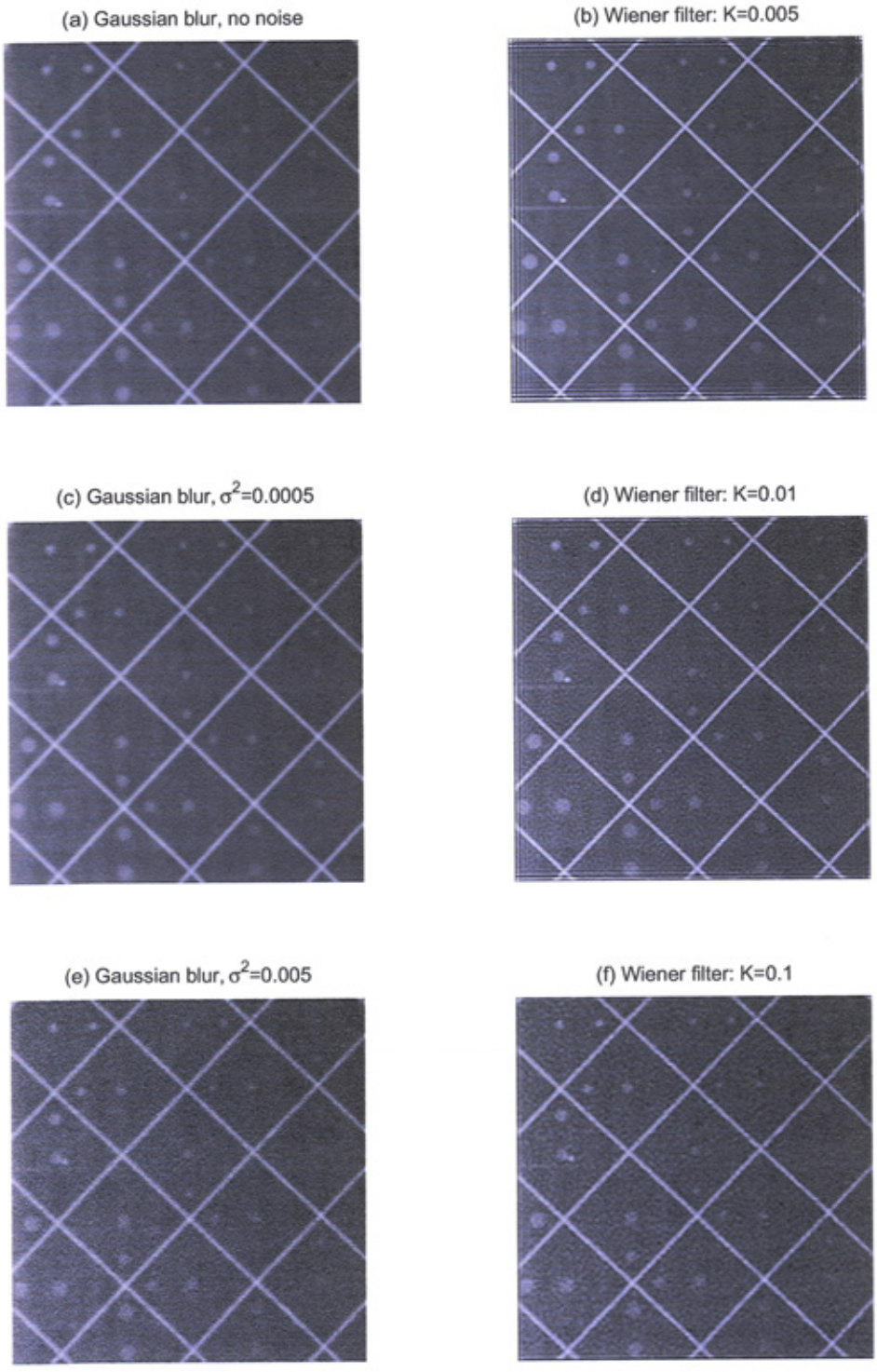

Figure 2.8: Deconvolution of the Gaussian blurred image with increasing noise levels using Wiener filtering.

and the more realistic images presented in subsequent chapters suggest that the de-blurring effectiveness of the Wiener filter becomes more limited as the SNR gets worse. The lack of literature (beyond that of a few specialised imaging applications) serves to emphasise that the general usefulness of this technique in improving noisy radiological images is regarded as somewhat limited.

2.3.5 Regularised deconvolution

It was shown in section 2.3.1 that the ill-posed nature of the restoration problem means that there may be many (or infinitely many) solutions $\hat{\mathbf{x}}$ that are consistent with the data $\mathbf{d}$. Some scheme of selecting the appropriate solution from the ensemble of possible solutions needs to be employed. This is usually achieved by defining *a priori* what the 'best' solution should look like. One such definition is that optimality of the restored image be based on a measure of smoothness (for example, the Laplacian of the image). It is also sensible to introduce constraints on the possible solutions; these may take the form of ensuring positivity is enforced in the solution (as negative pixel values are nonsensical in the image formation model), or that the total flux of the restored image is the same as the degraded image.

Let F be a function of the image pixels and C_i be constraints on the solution. The deconvolution strategy here is to optimise the function F subject to the constraints being met. This achieved by forming an objective function Q from these using some as yet undetermined Lagrange multipliers α_i:

$$Q = \mathrm{F} + \alpha_i C_i \tag{2.26}$$

The numerical procedure is then find a solution $\hat{\mathbf{x}}$ which optimises this objective function. This process is called regularised deconvolution. The regularising function, F, prevents the solution fitting the data (and the noise) too closely. The magnitude of α, which gives more or less weight to the constraints C_i, is adjusted to allow some definition of 'restoration' to be satisfied. There are many regularised deconvolution methods available and we discuss some of the most widely used next.

Constrained least-squares

A least-squares restoration is one in which the mean squared error between mock-data (based on an estimate of the solution) and the observed data is minimised, i.e.

$$\|\mathbf{e}\|^2 = \|\mathbf{d} - A\hat{\mathbf{x}}\|^2 \tag{2.27}$$

Least-squared restorations of noisy images tend to converge to restorations which fit the data too strongly: noise in the original data is strongly evident, usually as a graininess, in the restored image. An alternative is to use constrained least-squares restoration. The smoothness of the restored image can be quantified by its Laplacian (a gradient function: a measure of how quickly the pixel luminance within the image is changing with position). A restoration process which minimises the Laplacian of the estimate, subject to the constraint of equation 2.27 does rather better and will impose a degree of smoothness on the solution. Constrained least-squares techniques can be implemented in both the spatial and frequency domains. The frequency version of the filter is:

$$\hat{X}_{uv} = \left[\frac{1}{H_{uv}} \frac{|H_{uv}|^2}{|H_{uv}|^2 + \alpha |P_{uv}|^2} \right] D_{uv} \tag{2.28}$$

where equation 2.28 is derived by Gonzalez and Woods (1992). An iterative procedure to determine the appropriate value of α based on an estimate of the noise variance is also detailed. P_{uv} is the Fourier transform of the Laplacian, which in the spatial domain is simply the convolution function:

$$p_{ij} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{pmatrix} \tag{2.29}$$

Constrained least-squares has been applied to deconvolution of scintigrams (Boardman 1979)** when Wiener filtering of such images was unsatisfactory. Sutton and Kempi (1992) applied the method to deconvolve data obtained from renal scans and noted the superiority of the constrained least-squares over conventional Fourier inversion techniques.

Maximum Entropy deconvolution

The regularising function used to impose smoothness on the solution is the image entropy. The entropy of an image is a measure of structure in the information content. The entropy of the restored image is a suitable regularising function for some image formation models; an image solution is sought which has the least uncertain structure of all the solutions that are consistent with the measured data. This is equivalent to choosing the solution which has largest global entropy measure, i.e. the image having maximum entropy. RRegularised deconvolution using the entropy function requires optimization of an objective function of the form:

**See also page 82

$$Q = S + \alpha\chi^2 \tag{2.30}$$

where S is entropy in the reconstructed image and χ^2 is a statistical measure of the misfit between the observed image and data based on the estimate of the 'true' image. The role of a proper definition of χ^2 will be given in chapter 3. Some of the commonly used forms of S are listed below (Cornwell 1982):

$$S = -\Sigma_i x_i \ln(x_i) \tag{2.31a}$$
$$S = -\Sigma_i x_i [\ln(x_i/b_i)] \tag{2.31b}$$
$$S = -\Sigma_i \ln(x_i) \tag{2.31c}$$
$$S = -\Sigma_i \sqrt{\ln(x_i)} \tag{2.31d}$$

The values x_i represent pixel intensities in the restored image. The values b_i in equation (2.31b) represent default model values - pixel values the restored image should take in the absence of data. A more complete discussion of these default image model values is postponed until Chapter 3.

The choice of regularising function will depend on the situation, with some leading to better restorations than others. There might also be good theoretical or pragmatic reasons for choosing one function over another: see, for example, Wu (1997) and our arguments in Chapter 3. It has also been argued (Titterington 1985) that entropy should be regarded as a pragmatic regularising function with sensible smoothing properties independent of any theoretical justification for its use. However, in the following chapter we argue that a scheme of regularised deconvolution using the entropy function is appropriate for the medical imaging problems discussed in chapter 1.

2.3.6 Related deconvolution methods

Although we have restricted our attention to just a few methods in this chapter it is worth mentioning some notable absences with explanations of why they were not considered as suitable techniques for radiological image processing.

Basic iterative deconvolution

Iterative techniques, operating in either the spatial or frequency domains, can be quick and effective tools when applied to some image data. Consider the simple technique known as basic iterative deconvolution applied in the spatial domain to an observed image d_{ij}. The method begins by forming

an estimate $\hat{x}^{(0)}$ of the 'true' object distribution. This estimate is blurred with the known system PSF, denoted by h_{ij}, to form a 'trial image'. If the trial image is equal to, or very close to the observed image then the method terminated because we have found a solution. If there is a difference between the two, a small fraction of the residuals are added to the initial guess to obtain $\hat{x}^{(1)}$. The method continues iteratively:

$$\hat{x}_{ij}^{(n+1)} = \hat{x}_{ij}^{(n)} + \alpha(h_{ij} * \hat{x}_{ij}^{(n)} - d_{ij}) \tag{2.32}$$

Basic iterative deconvolution derives from the method of Van Cittert (Van Cittert 1931; Jansson 1970; Thomas 1981) which was originally used to resolve spectral lines. Despite the advantages in terms of speed and ease of implementation these methods are not particularly robust when applied to noisy images. In such cases the parameter α must be made so small as to reduce progress to a crawl. Convergence is rarely guaranteed in practice with most implementations controlling the algorithm according to the number of iterations.

CLEAN deconvolution

CLEAN deconvolution was developed by astronomers from a need to process aperture synthesis data from radio interferometers (Högbom 1974; Schwarz 1978; Clark 1980). CLEAN is a form of subtractive deconvolution (Bates and McDonnell 1989). The CLEAN algorithm iteratively builds a CLEANed image by adding scaled components to an initially empty image and results in a radio map devoid of unwanted sidelobes. The original method was modified by Keel (1991) to handle optical images and is known as σ-CLEAN.

CLEAN and variants of it are based on the assumption that the true image is well represented by point sources in an otherwise empty image. While this assumption can be valid for astronomical objects such as star clusters, or distant galaxies and quasars it is generally not the case for radiological images, which can contain detail and structure on many scales. Our experience of applying CLEAN to test images in this work shows that any extended structures are reconstructed with a rather grid-like appearance. CLEAN tends to give a varying spatial resolution in the reconstructed image because areas with a high signal-to-noise ratio are preferentially restored. The restored image can have regions where little or no reconstruction has taken place.

Richardson-Lucy deconvolution

The method due to Richardson (1972) and Lucy (1974) is to be found in widespread use in the astronomical community. Richardson-Lucy (RL) deconvolution was used on optical images from the Hubble Space Telescope before corrective optics were installed. RL deconvolution can be implemented in the spatial domain using convolutions or in the frequency domain using Fourier transformed variables (as is the case in equation 2.33). The procedure is iterative and the next iterate is calculated from:

$$\hat{X}_{uv}^{(n+1)} = \hat{X}_{uv}^{(n)} \left[H_{vu} \frac{D_{uv}}{H_{uv}\hat{X}_{uv}^{(n)}} \right] \tag{2.33}$$

Solutions found by RL deconvolution are positive and the total flux of the original image is also conserved. RL deconvolution converges to the maximum likelihood solution for Poisson statistics in the data (Shepp and Vardi 1982). Noise amplification is a problem with this method and the pragmatic approach usually adopted is to stop the iterations before the restored image becomes dominated by noise. In other words, although RL deconvolution converges to the mathematically correct solution, a better quality restoration is usually arrived at before the method converges (Bi and Boerner 1994). RL deconvolution has a tendency to overfit the restored image to the data (and therefore, the noise). A more sophisticated version called Damped Richardson-Lucy (White 1994) was devised to address the noise amplification problems of the conventional method.

In most practical implementations of RL deconvolution the termination criterion of the method is defined on the number of iterations. As with the Wiener filter, the ideal stopping point is usually obtained by trial and error. RL deconvolution represents a compromise between the quicker (but sometimes unsatisfactory) method of Wiener filtering and the more computationally expensive (but usually superior) regularised methods such as Maximum Entropy deconvolution.

2.3.7 Comparison of restoration methods

Some of the restoration methods discussed in this chapter were tested on blurred and noisy images. Test images were generated by blurring (with a known PSF) and adding noise (of a known standard deviation) to an image considered to be the 'true image'.

Method	Blur only	Blur $\sigma^2 = 10^{-4}$	Blur $\sigma^2 = 10^{-3}$	Blur $\sigma^2 = 10^{-2}$	Average time (s)
Inverse filter	—	—	—	—	—
Pseudoinverse filter	19.5	3.1	0.9	0.8	1.4
Wiener filter	19.8	17.6	2.1	1.9	2.1
Basic iterative deconvolution	19.5	12.8	1.5	0.9	24.0†
Constrained least-squares	17.6	11.5	2.0	1.4	20.1
Richardson-Lucy	21.1	18.8	3.5	2.5	27.3†
Maximum Entropy	22.4	19.0	4.7	3.2	43.5

Table 2.1: SNR improvements obtained after applying restoration methods to a blurred, increasingly noisy test image. The last column shows the time (averaged over the 4 trials) to arrive at a solution († terminated after 10 iterations). All methods were implemented with Matlab.

Quantifying success of restoration tests

When an image is processed with a restoration procedure it is sometimes obvious to the eye that an improvement in 'image quality' has taken place. However, it is necessary to define a way of quantifying the amount of improvement. Several measures exist and were discussed in Chapter 1. For this experiment a suitable measure of the restored image quality is the improvement to the SNR between the original (SNR_O) and restored image (SNR_R). We define this as a ratio:

$$\text{SNR improvement} = \frac{\mathrm{SNR}_R}{\mathrm{SNR}_O} = \frac{\mathrm{S}_R/\sigma_R}{\mathrm{S}_O/\sigma_O} = \frac{\mathrm{S}_R\sigma_O}{\mathrm{S}_O\sigma_R} \tag{2.34}$$

The quantities S_O and S_R represent signal strengths (in terms of pixel luminance values). The numbers σ_O and σ_R are measures of the standard deviation of the noise. Signal content can be estimated from features in the image and noise levels can be estimated from regions containing zero (or constant) signal.

Our true image is of a region of the CDMAM phantom (see section 2.3.2). Examples of restorations have already been used to illustrate some the methods but we show the results of more rigorous and complete restoration experiments here. The SNR improvements over the original images for each method are given in Table 2.1.

Results

Table 2.1 shows the SNR improvement ratios for images processed with the restoration techniques listed in the first column. Scores with a ratio greater

than unity show that the SNR of the restored image was greater than that of the original. Conversely, two of the scores are less than unity - the restoration process made the image worse!

As was already seen in section 2.3.3, the inverse filter completely fails to recover the image, even when no further noise has been added to the image. The pseudoinverse filter does reasonably well in the no-noise case when an appropriate value of ε is found but is far less successful at restoring in the presence of medium and high noise levels. The Wiener filter fares rather better than the pseudoinverse and some improvement in image quality is seen in all the restorations. The time taken to search for optimal parameters for these methods is not reflected in the above table. Van-Cittert deconvolution reconstructs the image well in the absence of noise but the restorations become very poor as the noise levels increase. Constrained least-squares does even better. The final two methods - Richardson-Lucy and Maximum Entropy are the most successful and seem (according the figures in the table) to have just a marginal difference between them in restored image quality.

The SNR table does not tell the full story. Close inspection of the results obtained with constrained least-squares restoration, Richardson-Lucy deconvolution and maximum entropy are presented in figure 2.9. Table 2.1 suggests the restoration quality of Richardson-Lucy is comparable with Maximum Entropy, with the computational expense of the latter making it a barely worthwhile endeavour. However, the images themselves show the de-blurring effect of Maximum Entropy to be superior to that of Richardson-Lucy.

2.4 Discussion

With the obvious failure of direct Fourier inversion techniques and the limitations of Wiener filtering it seems to be accepted among radiologists that image de-blurring is of little value. This opinion is supported by a lack of substantial literature and the absence of commercial use of such techniques.

However, there exist methods of regularised deconvolution which are more robust when dealing with noisy, blurred images. In particular maximum entropy deconvolution would seem to be a potentially fruitful but largely unexplored technique in mammographic and tomographic imaging. The complexity of the numerical procedure needed to obtain a solution and the significantly slower processing time is countered by the superior restorations it produces. In 1978 it was noted by Gull and Daniell that the time to process a 128×128 image on an IBM 370/170 machine with maximum entropy was 3 minutes. A quick experiment using a same-sized image on a modern In-

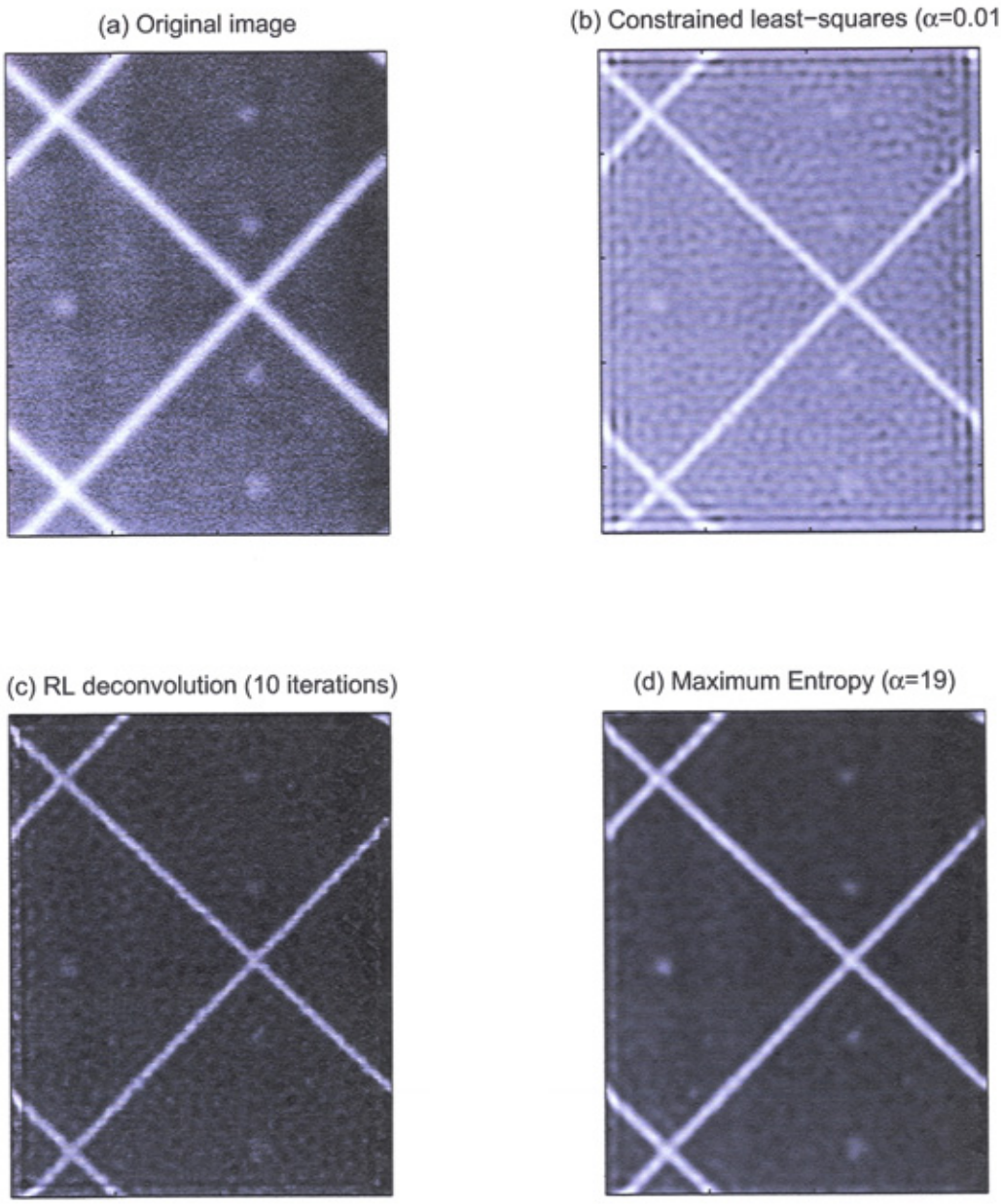

Figure 2.9: Comparison of restoration techniques. (a) is a section of the CDMAM phantom, artificially blurred and noise degraded. (b) Constrained least-squares deblurs the image at the expense of introducing a mottled background. (c) Richardson-Lucy restorations suffer the same problem. (d) Maximum entropy handles the noisy background in a much more satisfactory way and de-blurs features deemed to be 'real'

tel Xeon 3.2GHz machine arrives at a solution in less than 35 seconds.[††] Processing times have shortened not only with the availability of more computer processing power, but also with advances in how the maximum entropy algorithm is implemented. It is the opinion of the author that using maximum entropy to process images with *several million* pixels in a reasonable period of time on desktop PCs would not have been feasible until perhaps the mid-1990s.

Image processing methods in this chapter were presented as a choice between enhancement and restoration. In practice it may be useful or indeed necessary bring both methodologies to bear on a particular imaging problem.

[††]Matlab implementation. The commercial MEMSYS version does rather better and converged to a solution in just 18 seconds.

CHAPTER 3

Bayesian image restoration

In Chapter 2 it was stated that image restoration required the inversion of a matrix that was ill-conditioned or possibly singular. In this chapter we justify that statement and its consequences. A deconvolution strategy called Maximum Entropy Method (MEM), derived from Bayesian statistical inference, is offered as an appropriate way of estimating the 'true' image distribution from the noisy observed data. We introduce a Hybrid MEM scheme to handle noisy radiological images in a more satisfactory way by modifying the so-called 'Historic' method to introduce pixel correlations into the restoration.

3.1 Ill-posed inverse problems

In section 2.3 a simple model for a radiological imaging system was introduced. The equation for a linear image receptor with no system noise was introduced as:

$$d(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(i,\alpha,j,\beta)x(i,j)\,d\alpha d\beta \tag{3.1}$$

This is a Fredholm integral equation of the first kind and describes the forward map from hidden space to data space. We can simplify the notation of this mapping transformation to:

$$d = F\{x\} \tag{3.2}$$

Where:

$$F\{x\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(i,\alpha,j,\beta)x(i,j)\,d\alpha d\beta \tag{3.3}$$

Obtaining the solution x to the image restoration problem is an inverse problem which requires the determination of the transformation F^{-1} such that:

$$F^{-1}\{d\} = x \tag{3.4}$$

Hadamard (1902) proposed that a problem is ill-posed if, under appropriate conditions, the solution fails to satisfy one or more of the following statements:

1. The solution exists
2. The solution is unique
3. The solution is stable

If the inverse mapping F^{-1} does not exist the problem is said to be singular and a solution will also not exist. In this case the recovery of the true image distribution is not possible (although an approximate solution may be sought). Conditions leading to a singular problem are given in Appendix B.1. If the inverse mapping exists but is not unique then the solution will depend on which F^{-1} is chosen. Finally, if the inverse mapping exists and is unique it may not be stable. Instability in this case is described mathematically as:

$$F^{-1}\{d+\epsilon\} = x+\delta \tag{3.5}$$

where $\delta \gg \epsilon$. In words: a small change to the observed data, perhaps caused by inherent measurement errors, is amplified by the inverse mapping to a large change in true image distribution. It is shown in Appendix B.2 using the Riemann-Lebesgue lemma that the image restoration techniques based on equation (3.1) are ill-conditioned and unstable. If we also take into account the presence of noise (e.g. random sensor noise) in the imaging system the we rewrite equation 3.1 as:

$$d(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(i,\alpha,j,\beta)x(i,j)\,d\alpha d\beta + n(i,j) \tag{3.6}$$

In the idealised case of a linear relationship between the true object and recorded image where $n(i,j) = 0$ then a 1:1 unique mapping can exist between d and x (if no singularity exists). However, the presence of noise

means that a unique relationship between the two distributions can not exist (Andrews and Hunt 1977, page 116).

Therefore image restoration is an ill-posed problem because one or more of Hadamard's statements are not satisfied. Ill-posed problems are solved by introducing some degree of 'well-posedness' into the formulation. For example, if no solution exists (1) then we redefine the problem, perhaps by relaxing some constraints on allowable solutions. Such action may result in many or infinitely many solutions being consistent with the measured data, which is a violation of (2). Some scheme for selecting one solution from all feasible solutions must then be followed. Issues of stability (3) can be investigated once the appropriate solution has been determined.

In this chapter we shall justify our choice of appropriate solution as being the one with the maximum entropy.

3.2 Bayesian methods of image processing

There is a large body of literature describing methods which aim to improve the quality of image information content through restoration techniques. Medical images are usually measurements of photon flux, and the data is usually noisy and often incomplete; statistical methods have yielded some robust methods of estimating the 'true' image distribution in these circumstances. Snyder et al. (1992) describe a method which maximises the mean value of the log-likelihood for quantum noise limited data. It has been shown that this is equivalent to minimizing Csiszàr's I-divergence (Csiszàr 1991), a quantity equal to the negative of the entropy expression, $-S$, given in equation (3.32). The usefulness of Bayesian restoration stems from the fact that it allows the incorporation of sophisticated *a priori* knowledge into the formulation of the restoration method, while quite naturally enforcing desirable properties such as positivity in the restoration. It has been argued by Skilling (1989) that in the absence of further prior information entropy is the only consistent prior for positive, additive images. O'Sullivan et al. (1998) give a summary of these methods in terms of information theoretic image formation models.

3.3 Principles of Maximum Entropy

3.3.1 Definition of entropy

The concept of entropy was introduced into thermodynamics in the middle of the 19th century by Clausius and the entropy expression for statistical mechanics was derived by Boltzmann. The statistical mechanics definition of entropy is a measure of the disorder within a thermodynamical system. The entropy of an image is a property which comes from information theory, established by Shannon (1948). Shannon showed that the uncertainty of a random variable taking a discrete and finite number of values $x_i \ldots x_n$ could be quantified in terms of its probability distribution:

$$H = -\sum_{i}^{n} p_i \ln p_i \tag{3.7}$$

where p_i is the probability associated with each x_i. Shannon's entropy (H) is sometimes called information theoretic entropy to distinguish it from the form used in thermodynamical situations. In the remainder of this work it is to be understood that references to 'entropy' refer to the information theoretic form.

3.3.2 Historic MEM

The links between statistical mechanics and information theory were established by Jaynes (1957a,b, 1968). Image restoration using MEM was first described in a landmark paper by Frieden (1972). The driving force behind practical implementations of the method came from radio astronomy and the need to improve radio maps of the sky (Gull and Daniell 1978; Shore and Johnson 1980). Those implementations of MEM evolved as an alternative to CLEAN for working with the large images obtained from radio telescope interferometers. For images with more than about 1 million pixels MEM proved to be a more efficient deconvolution method. For a more detailed review of how MEM was initially developed for use in image restoration problems and a justification of why such restorations should be treated with confidence (beyond the arguments given below) the reader is directed to see Jaynes (1982).

MEM is a deconvolution technique derived from the forward map (3.8) for the imaging system, which relates postulated hidden data x_{ij} to the observed data d_{ij}.

$$d_{ij} = x_{ij} * h + n_{ij} \tag{3.8}$$

Equation (3.8) is the forward map for the imaging system and is the same one that was discussed in section 2.3. The hidden image x is mapped to the observed data d through a convolution with a point-spread function h, which characterises the geometric blurring of the imaging system. The image is then further corrupted by additive noise n to produce the observed image d.

The goal of image restoration is to obtain a solution $\hat{x}$, which approximates to the hidden image x, as closely as the data and noise allow. We outline briefly the Historic MEM approach to the problem here, but for a more complete review see Skilling and Gull (1985). A trial restoration $\hat{x}$ is obtained and used as an initial guess for the hidden image x. The trial restoration is blurred, using the PSF of the forward map, to generate mock data $\hat{d}$:

$$\hat{d}_{ij} = \hat{x}_{ij} * h \tag{3.9}$$

The notation in the following analysis has been simplified in relation to equation (3.8), in that the data etc. are represented as vectors rather than 2D arrays; equation 3.9 becomes

$$\hat{d}_i = \sum_j h_{j-i}\, x_j \tag{3.10}$$

The χ^2 goodness of fit statistic is used to measure the degree of misfit between the observed data and mock data:

$$\chi^2 = \frac{\sum_i (d_i - \hat{d}_i)^2}{\sigma^2} \tag{3.11}$$

where σ^2 is the variance in the noise, here taken to be constant across the image. This approach is a simplification which we acknowledge in Appendix C.1. For some images, such as the skulls in chapter 6, the variation in noise levels is incorporated via equation C.1.

It might be thought that a good approach would be to minimise the degree of misfit χ^2 by choosing a suitable $\hat{x}$, but this process is equivalent to the straightforward matrix inversion:

$$\hat{x} = A^{-1} d \tag{3.12}$$

where A is the matrix representation of the linear forward map (Equation 3.8). Proof of this is given in Appendix C.1. As we have already seen, the matrix A is ill-conditioned, i.e. almost singular, so for a given d there are

many vectors $\hat{x}$ which almost satisfy equation (3.12), not necessarily close together. Thus if d contains even a small amount of noise, the resulting $\hat{x}$ can be far from the true image x; in other words the reconstruction is then dominated by noise rather than data, often referred to as over-restoration. In the presence of noise we would not in any case expect χ^2 to be minimised, but rather to be reduced to the appropriate value $\chi^2 = N$, where N is the number of pixels in the image. Even when A is not ill-conditioned there are many $\hat{x}$ which satisfy this criterion. The problem is now the appropriate choice of $\hat{x}$.

Images as probability distribution functions

An image x_i is a sequence of positive numbers, which represent the number of photons detected by the image receptor. Firstly we can write:

$$M = \sum_i x_i \tag{3.13}$$

where M is the total luminance in the image. The image data can be normalised and represented by a sequence of proportions:

$$p_i = x_i/M \Rightarrow \sum_i p_i = 1 \tag{3.14}$$

Light intensities, represented as discrete luminance values are positive and additive between pixels in the image. Therefore:

$$p_{i \cup j \cup \ldots} = p_i + p_j + \ldots \quad i \neq j \neq \ldots \tag{3.15}$$

$$p_i \geq 0 \tag{3.16}$$

Equations (3.14), (3.15) and (3.16) satisfy the Kolmogorov axioms of probability theory. Thus, an image can be thought of as being a probability distribution function (PDF) and can be analyzed with the rules of probability calculus and Bayesian inference. Our image is now represented as:

$$X : \begin{pmatrix} x_1 & x_2 & \ldots & x_N \\ p_1 & p_2 & \ldots & p_N \end{pmatrix} \tag{3.17}$$

where p_i is the prior probability that the random variable X assumes the value x_i.

The entropic prior

Historic MEM treats the restoration process as a statistical inference problem based on Bayes' theorem and the aim is to obtain the most probable image $\hat{x}$ given the data:

$$P(\hat{x} \mid d) \propto P(d \mid \hat{x}) \times P(\hat{x}) \tag{3.18}$$

where $P(\cdot)$ represents the conditional probability. This equation is worth stating in words:

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior} \tag{3.19}$$

The likelihood $P(d \mid \hat{x})$ is determined from our knowledge of the forward map (equation 3.8); the image noise is assumed to be mainly quantum (photon) noise, which is modeled as a Gaussian process*. The likelihood term is quantified by the χ^2 distribution:

$$P(d \mid \hat{x}) \propto \exp(-\chi^2/2) \tag{3.20}$$

The choice of prior probability is a controversial aspect of MEM. We present the 'monkey' argument of Jaynes (1986) in support of our choice.

Consider an experiment in which each of the luminance quanta, of which there are M, are scattered randomly† over an initially blank image which is divided into N pixels. Each luminance quantum has an *a priori* chance of being in any pixel. At the end of the experiment we observe that x_i quanta have arrived in pixel i. The number of ways of arriving at a configuration $x_1 x_2 \ldots x_N$ is simply the combinatorial expression

$$\Omega = \frac{M!}{\prod_{i=1}^{N} x_i!} \tag{3.21}$$

where Ω is the number of permutations of all the luminance quanta, divided by the number of equivalent reorderings. The number Ω is referred to as the degeneracy of the scene. Taking logs of both sides of equation (3.21) we obtain:

$$\ln \Omega = \ln M! - \sum_i \ln(x_i!) \tag{3.22}$$

*We are approximating the Poisson distribution with a Gaussian distribution; this follows from the large numbers of photons involved in the image formation process (see page 16).

†By the traditional team of monkeys!

Since both M and x_i are large we can make use of Stirling's approximation for factorials[‡] with appropriate use of equation (3.13):

$$\begin{aligned}
\ln \Omega &\approx M \ln M - M - \left(\sum_i x_i \ln x_i - \sum_i x_i \right) \\
&\approx M \ln M - M - \sum_i x_i \ln x_i + M \\
&\approx M \ln M - \sum_i x_i \ln x_i \\
\ln \Omega &\approx -\sum_i x_i \ln \frac{x_i}{M} \qquad (3.23)
\end{aligned}$$

If we scale the luminance quanta in accordance with equation (3.14) we can treat it as a PDF. Substituting $x_i = Mp_i$ into equation (3.23) gives:

$$\begin{aligned}
\ln \Omega &\approx -\sum_i (Mp_i) \ln \frac{Mp_i}{M} \\
&\approx -M \sum_i p_i \ln p_i \\
&\approx MH \qquad (3.24)
\end{aligned}$$

where H is Shannon's information theoretic entropy seen earlier, in equation (3.7). We arrived at equation 3.23 by assuming that each luminance quanta had the same a priori chance of being in a particular pixel. If instead there we have some more *a priori* knowledge (embodied as a probability associated with each pixel) then the same reasoning leads to a modified expression of entropy given by:

$$\begin{aligned}
MS &= -\sum_i x_i \ln \frac{x_i}{Mp_i} \\
&= -\sum_i x_i \ln \left(\frac{x_i}{m_i} \right) \qquad (3.25)
\end{aligned}$$

where S is a modification of the Shannon entropy of equation (3.7). The $m_i = Mp_i$ would be observed as mean values of x_i. The entropic prior $P(\hat{x})$ is therefore:

[‡] $\ln F! \approx F \ln F - F$

$$P(\hat{x}) \propto \exp(MS) \tag{3.26}$$

where S is the entropy defined in equation (3.25). With this equation and equation (3.20) we can now specify the posterior probability given in equation (3.18)

$$\begin{aligned} P(\hat{x} \mid d) &\propto \exp(-\chi^2/2)\exp(MS) \\ &\propto \exp(MS - \chi^2/2) \end{aligned} \tag{3.27}$$

Values of x_i which maximise this probability must be sought. Degenerate inversions are likely to be able to achieve unrealistically small values of χ^2. An adjustable parameter, denoted by α, to bring χ^2 into its expected narrow statistical range is required. It does not matter to which term we attach α and for consistency with the Hybrid method discussed later, we absorb the factor M into it and attach it to the entropy term. Therefore a pragmatic modification of equation 3.27 is given by:

$$P(\hat{x} \mid d) \propto \exp(\alpha S - \chi^2) \tag{3.28}$$

Maximization of equation (3.25) subject to the constraint that $\Sigma_i x_i = M$ is a constant is equivalent to the unconstrained maximization of the entropy function:

$$S = -\sum_{i=1}^{N} x_i \left[\ln\left(\frac{x_i}{m_i}\right) - 1\right] \tag{3.29}$$

In the absence of data constraints the entropy is maximised when $\frac{\partial S}{\partial x_i} = 0$. Differentiating equation (3.25) gives the gradient of the entropy:

$$\frac{\partial S}{\partial x_i} = \ln m_i - \ln x_i \tag{3.30}$$

If we set this to zero we see that $x_i = m_i$ at the stationary point. Differentiating equation (3.30) gives $\partial^2 S/\partial x_i x_j = -\delta_{ij}/x_i$ confirming that this is indeed a maximum; S is maximised when $\hat{x}_i = m$, giving a flat featureless reconstruction. The numbers m_i are therefore default values to which the restored image pixels x_i will be set unless the data demands otherwise. The m_i represent a default model which can encode any *a priori* beliefs about how the restored image should look. We discuss further the choice of default model in section 3.4.4. We note for now that S is maximised when $\hat{x}_i = m_i$.

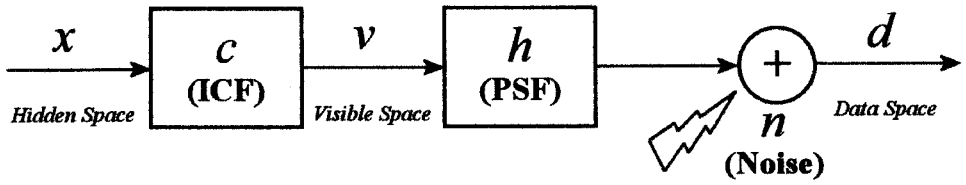

Figure 3.1: The image formation model for MEM. The hidden data is mapped to the visible space by the ICF, which in turn is blurred with the imaging system PSF and further degraded with additive noise, n, to produce the observed image data d.

3.4 Hybrid MEM

Images restored by Historic MEM tend to be grainy in appearance at the very finest scales: there can often be very little correlation between neighbouring pixels within the image. One of the axioms of MEM is that pixel correlations should not be introduced into the restored image, that is to say, restoration in one part of an image should not affect restoration in another part (Skilling 1988). However, such pixel correlations more often than not exist in the real world and usually between neighbouring pixels. We introduce a 'hybrid' formulation of Historic MEM which allows pixel correlations in the restored image. The image formation model is slightly different from the Historic MEM approach and is shown in figure 3.4. The corresponding forward map for the imaging system is:

$$d_{ij} = x_{ij} * c * h + n_{ij} = v_{ij} * h + n_{ij} \tag{3.31}$$

The Intrinsic Correlation Function (ICF), denoted by c, maps the hidden image x to the visible image v and is a way of encoding any expectations about pixel correlations into the forward map (see section 3.4.1). The visible image v is mapped to the observed data d through a convolution with a point-spread function h, which characterises the geometric blurring of the imaging system. The image is then further corrupted by additive noise n to produce the observed image d.

The restoration process is again treated as a statistical inference problem based on Bayes' theorem and the aim is to obtain the most probable image $\hat{x}$ given the data. The analysis proceeds exactly as for Historic MEM, the only difference being the form of the entropic prior:

$$S(\hat{x}) = \sum_i [\hat{x}_i - m_i - \hat{x}_i \ln(\hat{x}_i/m_i)] \tag{3.32}$$

This form is suitable for positive additive distributions which do not sum to unity and its derivation is discussed by Skilling (1988). Comparison with our previous definition of entropy (equation 3.29) shows that there is an additional term m_i present. There is no difference to the maximization process: it still corresponds to $x_i = m_i$ in the absence of constraints. However, the value of S at the maximum is different. Under Historic MEM the entropy had a value of $S = \Sigma_i m_i$ - a quantity which varied with the default image model being used. Skilling's entropy (equation 3.32) is also maximised at $x_i = m_i$ but the maximum value of the entropy is absolute, with $S = 0$ whatever default image model is used.

The rest of the formulation follows in much the same way as Historic MEM. A numerical scheme is required to find the maximum of the objective function:

$$Q = \alpha S - \chi^2 \tag{3.33}$$

The forward map now demands the mock data to be generated with $\hat{d} = \hat{x} * c * h$. The goodness-of-fit with the observed data is quantified by the χ^2 statistic given in equation (3.11). The maximisation of equation (3.33) assigns values to the hidden distribution $\hat{x}$. The final step is to convolve the solution with the ICF to obtain the restored 'visible' image.

3.4.1 Intrinsic Correlation Function

As has already been stated the ICF is a way of encoding *a priori* expectations on pixel correlations into the formulation of MEM. All the expectation of correlations are assigned to c so that x itself is *a priori* uncorrelated. This is achieved by defining the restored 'visible' image as a blurred version of an underlying hidden distribution:

$$v_{ij} = x_{ij} * c \tag{3.34}$$

In this way we can legitimately assign the entropic prior to x and the corresponding analysis and justification for MEM remains valid. The introduction of the ICF allows quite sophisticated incorporation of prior knowledge into the restoration. For example, c and x can be digitised at different resolutions.

For the image restorations in this work we use a fairly conservative scheme for enforcing correlations between neighbouring pixels: The ICF was modeled as a 3×3 Gaussian function relating a pixel only to its immediate neighbours:

$$c = \begin{pmatrix} 0.0751 & 0.1238 & 0.0751 \\ 0.1238 & 0.2042 & 0.1238 \\ 0.0751 & 0.1238 & 0.0751 \end{pmatrix} \tag{3.35}$$

In our tests of Hybrid MEM a larger ICF resulted in stronger correlations between more distant pixels; restored images looked very smooth with a poor spatial resolution. With the ICF given in equation 3.35, improvements in spatial resolution are not lost by our Hybrid MEM's enforcement of pixel correlations in the restored image. Figure 3.2 shows a comparison between Historic MEM and our Hybrid MEM in the restoration of a blurred and noisy image. The effect of applying an unsuitably large ICF (i.e. over enforcing pixel correlations) is also shown.

3.4.2 Implementation of MEM

Numerical methods

MEM can be thought of as a type of regularised deconvolution. It relies on a scheme of iterated forward maps rather than attempting to find a direct solution of the inverse problem. We describe briefly some numerical procedures to obtain the maximum entropy solution.

Gull and Daniell (1978) obtained the Historic MEM image using the *integral equation* approach of maximising the objective function for a fixed α

$$Q_{hist} = S_{\text{hist}} - \alpha\chi^2 \tag{3.36}$$

where S_{hist} was defined in equation 3.29. This leads to iterative solutions obtained with:

$$p_i^{(n+1)} = m_i \exp\left[-\alpha \frac{\partial \chi^2(p^{(n)})}{\partial p_i^{(n)}}\right] \tag{3.37}$$

where n denotes the nth iterate. The data and reconstructions are scaled to sum to unity - we are dealing with PDFs and the arguments for obtaining maximum entropy are valid (see page 68). This simple scheme is attractive because the exponential function ensures successive iterates are positive and it allows large values of p_i to develop quickly. In medical imaging applications (and in its original astronomical context) this is important because images

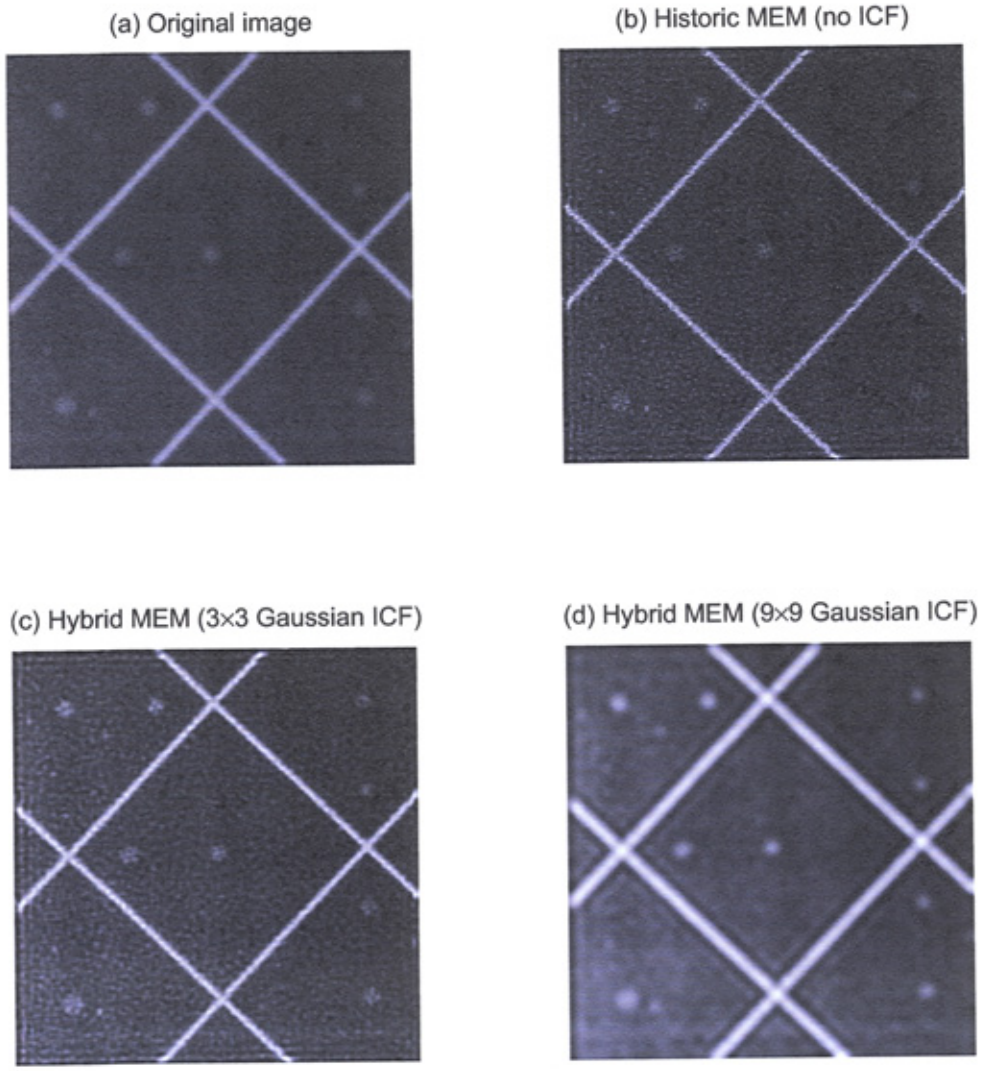

Figure 3.2: The effect of changing the size of the ICF on the restored image. (a) is an blurred, noisy image of part of the CDMAM phantom. (b) shows the Historic MEM reconstruction, which does not use an ICF. The reconstructed image is very grainy because the entropy function does not impose correlations. (c) is our Hybrid MEM using a conservative 3×3 ICF. (d) shows a reconstruction using a larger ICF - spatial resolution is lost as the image is oversmoothed. A flat default model ($m = 0.001$) was used in each case.

can potentially have a huge dynamic range. However, equation 3.37 is very unstable and successive iterates had to be smoothed with:

$$p_i^{(n+1)} = (1-\beta)p_i^{(n)} + \beta m_i \exp\left[-\alpha \frac{\partial \chi^2(p^{(n)})}{\partial p_i^{(n)}}\right] \tag{3.38}$$

where the parameter β determines the degree of smoothing. At large values of α the value of β often needs to be reduced so much that the method effectively slows to a stop in search of the solution.

Other optimisation techniques have been applied to finding the maximum entropy solution. For example, finding the maximum of the objective function Q_{hist} through the method steepest ascents leads to the following iterative procedure:

$$p_i^{(n+1)} = p_i^{(n)} + \gamma \frac{\partial Q(p^{(n)})}{\partial p_i^{(n)}} \tag{3.39}$$

for an appropriate value of γ. The method fails because for whenever γ is sufficiently large for high values of p_i to develop there are also pixels with negative values of $\partial Q/\partial p_i$ at which values of p_i become negative. The entropy function is only defined for $p_i > 0$ so these values must be 'clipped' after each iteration and set to some small positive value. This instability can result in painfully slow progress towards the maximum.

The steepest ascent algorithm is improved by using the conjugate gradient technique (Fletcher and Reeves 1964) or variants of it: for examples see Gill et al. (1981). Our preliminary investigations of maximum entropy used conjugate gradient methods and were implemented using Matlab.

Hybrid MEM: numerical solution with Matlab

The solution to our Hybrid MEM technique is obtained by maximising the objective function:

$$Q = \alpha S - \chi^2 \tag{3.40}$$

In practice we use numerical techniques to minimise $\hat{Q} = -Q$ which is an equivalent problem. Initial feasibility tests were undertaken using the Matlab Optimization Toolbox (The Mathworks Ltd, Cambridge, UK), particularly the constrained nonlinear minimization function *fmincon* in its LargeScale configuration, with a positivity constraint on each $\hat{x}_i$. This function employs a preconditioned conjugate gradient (PCG) method (Coleman and Li 1994, 1996). PCG methods for nonlinear functions require information about the first and second derivatives of the objective function $\hat{Q}$ where:

$$\hat{Q} = -\alpha S + \chi^2 \tag{3.41}$$

where S is the entropy defined in equation (3.29) and χ^2 is the degree of misfit between mock data and observed data given in equation (3.11). The gradient of equation (3.41) is:

$$\frac{\partial \hat{Q}}{\partial x_k} = -\alpha \frac{\partial S}{\partial x_k} + \frac{\partial \chi^2}{\partial x_k} \tag{3.42}$$

$$= -\alpha(\ln m_i - \ln x_i) - \frac{2}{\sigma^2} \sum_i A^T_{i-k}(d_i - \hat{d}_i) \tag{3.43}$$

where A denotes the combined action of the ICF and PSF, related by $A = c * h$. The Hessian matrix is obtained by differentiating equation 3.42 with respect to x_j

$$\frac{\partial^2 \hat{Q}}{\partial x_j \, \partial x_k} = -\alpha \frac{\partial^2 S}{\partial x_j \, \partial x_k} + \frac{\partial^2 \chi^2}{\partial x_j \, \partial x_k} \tag{3.44}$$

$$= \alpha \frac{\delta_{jk}}{x_j} + \frac{2}{\sigma^2} \sum_i A_{j-i} A_{k-i} \tag{3.45}$$

A derivation of the first and second derivatives of the entropy and χ^2 terms is given in Appendix C.2. If the image has N^2 pixels then the Hessian matrix will have N^4 elements. Realistic medical images may have $N^2 \approx 2 \times 10^6$ pixels, requiring construction of a Hessian with $\approx 1.6 \times 10^{13}$ elements. The second derivative of the entropy term contributes only to the main diagonal of the Hessian with the remaining off-diagonal contributions coming from the second derivative of χ^2, which is much less sparse. The computational expense of constructing the full Hessian is therefore enormous even when exploiting its sparse nature. Early tests of our Matlab implementation proved to be incredibly slow and this approach is infeasible for all but the smallest of images.

The non-diagonal components, due to the second derivative of χ^2, can be neglected in situations where the PSF is well defined (Cornwell and Evans 1985). The second derivative of χ^2 is instead approximated with a scaled identity matrix:

$$\frac{\partial^2 \hat{Q}}{\partial x_j \partial x_k} \approx \alpha \frac{\delta_{jk}}{x_j} - 2q\mathbf{I}$$

Cornwell and Evans argue that in most cases this is tantamount to ignoring any sidelobes of the MTF. The value of q is relatively unimportant, but should represent the power in the main lobe of the MTF. Experiments with different values had very little detrimental effect on convergence unless q was much less or much greater than unity. For the experiments in this work we simply used $q = 1$.

Numerical procedure

The procedure to obtain the maximum entropy solution from our Hybrid MEM technique is shown in figure 3.3. The Matlab code for implementing our Hybrid MEM is listed in Appendix D.5. Our procedure has two loops. The inner one, in which α is fixed, iterates towards the minimum value of $\hat{Q}$ for that α, and generates a corresponding value $\chi^2(\alpha)$. The second loop iterates over α to minimise $\chi^2(\alpha) - N$, and is terminated when $\chi^2(\alpha)$ falls within the narrow statistical range $N \pm (2N)^{1/2}$ (Jackson and Ward 1981). Typically each loop requires 20 to 30 iterations, and the final value of α is of order 20. After our successful initial implementation commercially available C++ software was purchased (MEMSYS5, Maximum Entropy Data Consultants, Bury St Edmunds, UK). This employs essentially the same scheme, but uses highly optimised choices for search directions, as described in Skilling and Bryan (1984) and Gull and Skilling (1999). The MEMSYS5 software can handle larger images than our initial implementation, with much shorter processing times.

Noise estimation

MEM requires an estimation of noise levels within the image. In the test images of section 2.3.2 the noise variance was set at a pre-determined level before the restoration. In most practical situations it must be estimated from the observed data. Several test objects (including the Leeds TORMAM phantom) are used in Chapter 4. For those experiments the variance is calculated by finding a section of the image where the signal is assumed constant and then assuming that variations across a small region are caused by noise.

In real radiological images, such as those encountered in Chapter 6, there may not be any regions from which a useful 'flat' area can be chosen; an alternative procedure to estimate noise levels would be required. Analysis of methods to estimate the noise component of anatomical images using ideas from information theory have been carried out by Tapiovaara and Wagner (1993) and Bochud et al. (1999). A method of extracting estimates of noise

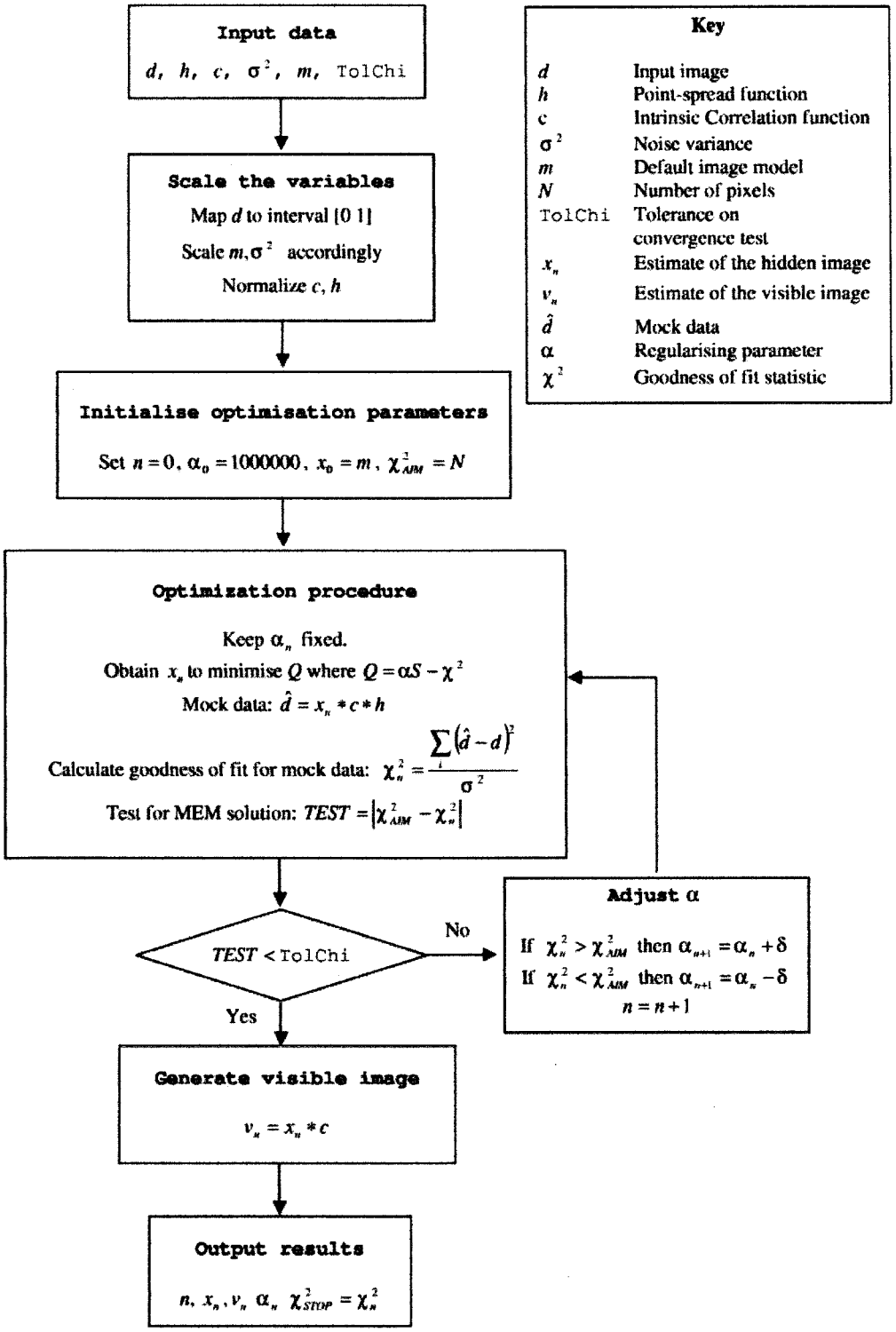

Figure 3.3: Flowchart showing the procedure for obtaining the hybrid maximum entropy reconstruction.

variance in MR images, which is also applicable to radiological images, is discussed by Sijbers et al. (1998). Another approach is to apply a wavelet transform to the image and estimate the noise using coefficients in the first scale. This method, using the non-subsampled biorthogonal discrete wavelet transform, was demonstrated by Costaridou et al. (2001) in a mammography setting. The estimation of noise variance within images obtained from various imaging systems was recently tackled using a statistical autoregressive method (Kamel and Sim 2004).

A fairly good estimate of the noise is important to the restoration process. If we underestimate the noise then calculated values of χ^2 will be too large and the algorithm will look for a solution too far from the maximum entropy solution, i.e. one that fits the data and noise too closely. Conversely if we overestimate the noise then χ^2 will be too small and the algorithm will find a solution too close to the default image model m, i.e. one that gives too much weight to the entropy and looks oversmoothed.

3.4.3 The effect of α on the restoration

To demonstrate the effect of α on the restored images some values of α were chosen manually. Restorations of the blurred, noisy images of the CDMAM phantom are shown in figure 3.4.

Inspection of equation (2.30) shows that a large value of α will bias pixel values of the reconstructed image towards values in default image model. Unconstrained maximisation of equation (3.32) gives a reconstruction equal to the default image model ($x_i = m_i$). Successively smaller values of α move the restored image away from the default image and more in line with the observed data. If α is made too small then the smoothing properties of the method are lost and the data (and noise) begin to dominate the reconstructed image. Setting $\alpha = 0$ results in restoration which has no regard to the entropy at all. In this case the MEM has been reduced to a least-squares minimisation procedure.

In the example presented in figure 3.4, the multiplier α, which determines the balance between over-restoration (more weight to the data and noise) and under-restoration (more entropy weighting), was chosen manually to give the best result.

3.4.4 The default model

The Historic and Hybrid MEM algorithms considered earlier initially use the default image model as an estimate of the restored image ($x_i^{(0)} = m_i$). Successive estimates of the true image distribution are gradually moved away

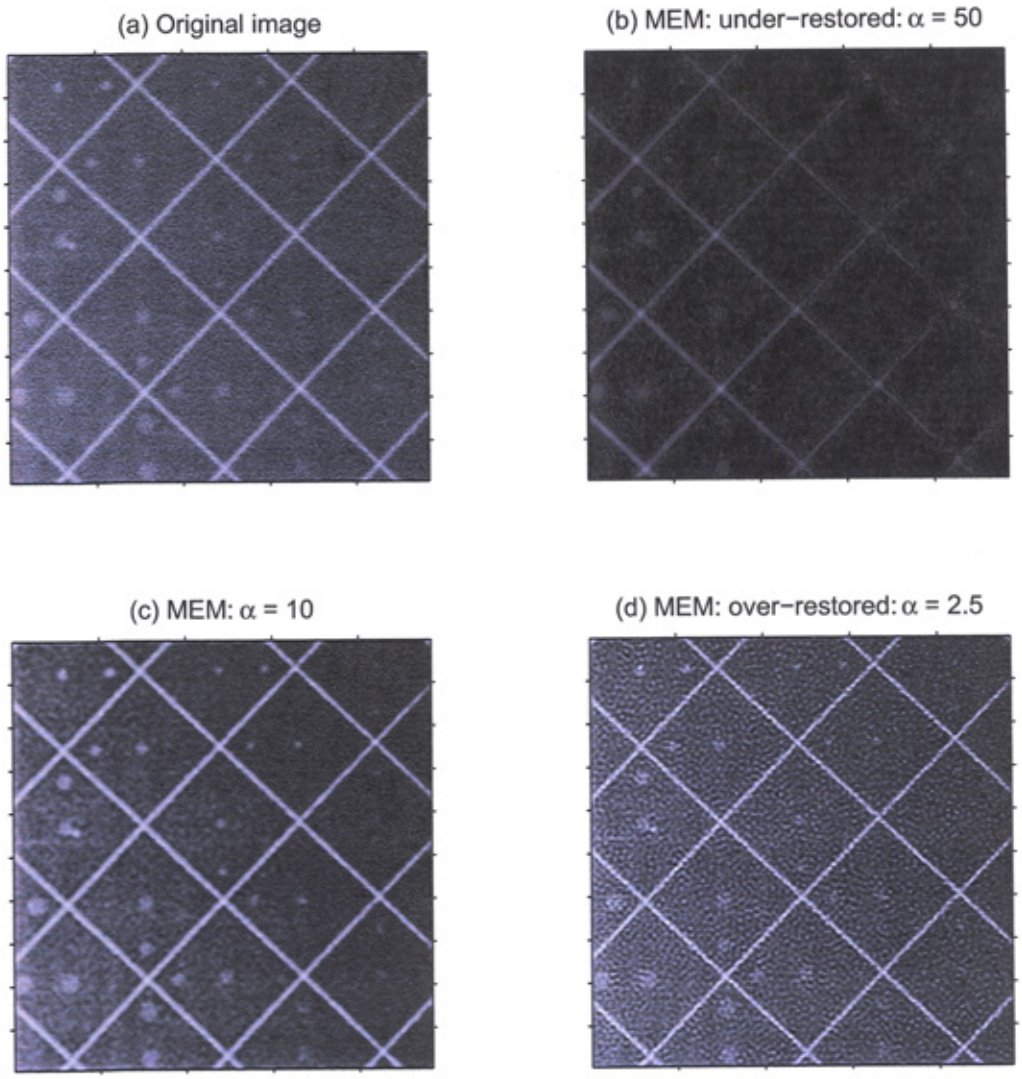

Figure 3.4: (a) is the original noisy, blurred image of the CDMAM phantom. A flat default model ($m = 0.001$) was used in each of the MEM restorations. (b) used a large value of α, giving too much weight to the entropy and resulting in an under restored image which has not strayed far from the default image model. (c) used a value of α chosen manually to balance smoothness against sharpness. (d) used a value of α which gave too little weight to the entropy term and which led to an over-restoration dominated by noise.

from the default model using numerical procedure of section 3.4.2. The eventual solution is as consistent with the default model image as the data and noise allow.

The default image models considered so far have been flat with low pixel values ($m_i = 0.001$ for $i = 1 \ldots N$); this is an appropriate choice if we have only minimal preconceptions. In the context of medical imaging this choice of m corresponds to claiming no prior knowledge of the image beyond that of low pixel luminance values in the absence of data (or put another way, very high photon counts if there is not sufficient evidence of absorption by bone or tissues).

If it is known that the solution has a 'preferred' form then the default levels m_i for pixel i can be chosen accordingly. A non-flat default level would be appropriate for example when the x-ray illumination is not uniform or when other knowledge of the restoration is known *a priori*.

3.5 MEM applications in medical imaging

Maximum entropy deconvolution has been applied to planar scintigrams by Simpson et al. (1995). Scintigrams are 2-D representations of a 3-D object through the use of radioisotopes. Scintigraphy differs from projection radiography in that blurring is caused primarily by Compton scatter (page 10) of emitted photons by patient tissues. Degradations caused by the finite resolution of the camera and image receptor are also present. An additional distortion associated with this technique is that structures nearer the radiation source are rendered more clearly than those further away. Their implementation was the Historic MEM described earlier, and using a smoothed version of the data as the default image model m. The authors tested the method on planar scintigrams of a Williams test object to evaluate image quality and demonstrated improvement on an image of bone. They argued that MEM could be used to reduce patient dose without sacrificing image quality in scintigraphy.

Historic MEM processing in Magnetic Resonance Imaging (MRI) was discussed by Moran (1991). The treatment of lumbar spine scans were presented as a demonstration of the usefulness of MEM.

3.6 Discussion

Image restoration of medical images based on the forward map described in this chapter is at best, an ill-posed problem and at worst a potentially singu-

lar problem with exact recovery of the true object distribution not possible. Relaxing constraints on allowable solutions requires some method for choosing a solution which is consistent with the measured data; we choose the one with maximum entropy. Historic maximum entropy has been applied to many imaging situations but solutions obtained from it can be very unsatisfactory, showing a graininess which is caused by the method's disregard for pre-existing correlations in the image model formulation. Our Hybrid version of the method, introduced and implemented in this chapter, overcomes some of the shortcomings of Historic MEM by introducing an Intrinsic Correlation Function to encode prior expectations of pixel correlations. Resulting images are smoother without sacrificing improvements to resolution.

Comparison with SVD

The MEM method has much in common with the SVD solution given by equation 2.17, but has different error and objective functions. The SVD method involves only elementary matrix manipulations, rather than an iterative technique, and in principle should be faster and easier to implement than MEM. In practice this is not the case as indicated on page 45. Practical implementation of SVD would probably be best achieved using a similar iterative approach.

CHAPTER 4

MEM and magnification mammography

Through a series of experiments using both specially constructed and clinically recognised test objects, the effectiveness of MEM applied to magnification mammography will be explored. The specific objectives of these experiments are as follows:

Conventional set-up to obtain images at conventional magnifications with a fine focal spot and use MEM to improve image quality.

Broad focal spot set-up where images are obtained at a conventional magnification and using a broad focal spot. These images will be unacceptably blurred and so this is not done in practice. MEM might be used to de-blur such images whilst maintaining or improving the SNR. Can we make these images as good as those taken under the conventional set-up?

Higher magnification The blurring associated with magnifications higher than 2.0 is so great that, even with a fine focal spot, they would not be used in practice. However, higher magnifications coupled with image de-blurring would be useful for clarifying the shapes of microcalcifications - a necessary step for better diagnosis of malignant or benign processes in the breast.

Dose reduction Reduction of the radiation dose to see how MEM copes with images which are much noisier than those obtained with typical

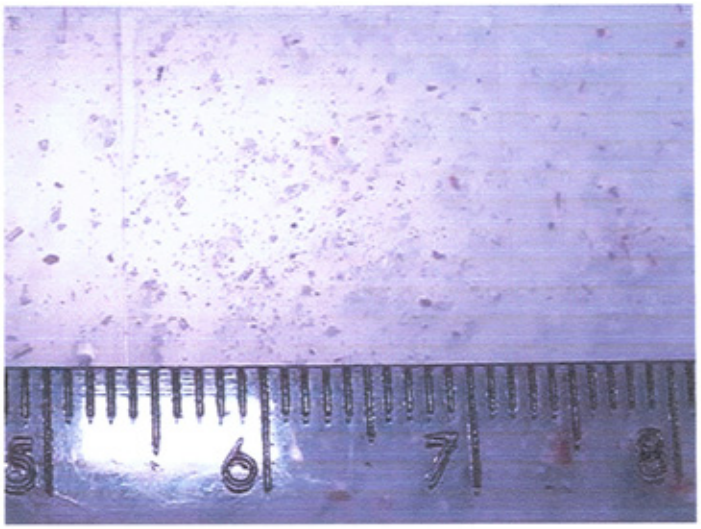

Figure 4.1: Photograph of part of the eggshell phantom. The eggshell phantom was designed to simulate microcalcifications in terms of their size and composition.

radiographic settings. Can MEM maintain image quality when noise levels are higher?

Initial experiments were carried out using the Matlab implementation of historic MEM (as outlined in Chapter 3) on small regions of simple test objects. Although software exists to implement MEM, it was desirable to judge the effectiveness of the method before purchasing the optimised commercial code. After successful early trials the later experiments used more realistic test objects and large images were processed with MEMSYS5, a commercially available implementation of MEM.

4.1 Initial experiments with MEM

Several test objects were constructed to evaluate the performance of MEM in reconstructing images with a simple structure. These test objects were made using sheets of perspex with approximate dimensions of 15cm×21cm and a thickness of roughly 0.5cm. Descriptions of these 'phantoms' and the structures they contained are given below:

Metal phantom Highly x-ray absorbing objects such as pins, a needle, washers and steel wool.

Bone phantom Small fragments and shavings of bone*. The composition of bone is primarily calcium phosphate (in the chemical arrangement termed calcium hydroxyapatite). Also present is the protein collagen.

Eggshell phantom Crushed eggshell fragments. The composition of eggshell is approximately 95% calcium carbonate with the remaining material containing magnesium, phosphorus and protein. Some microcalcifications are also composed of calcium carbonate (Fandos-Morea et al. 1988) so the properties of eggshell are very similar. The fragments ranged in size from 2mm to perhaps to $\approx$ 0.1mm. Part of this eggshell phantom is shown in figure 4.1.

A piece of brass foil with tiny 'pin pricked' pinholes was attached to each phantom so that a PSF could be obtained directly from the image of the particular phantom. Inclusion of the pinholes circumvented the need to obtain and rescale the PSF image in a separate step.

Notes on the measured PSF

The placement of a pinhole within the same field as the object also has the advantage of being a complete measure of the unsharpness present in the imaging system. The measured PSF from the pinhole in these situations is a combination of several sources of blurring:

1. Geometric unsharpness (caused by the finite size of the focal spot).
2. Blurring caused by the finite dimensions of the pinhole.
3. Receptor unsharpness is introduced when the latent image in the storage phosphor plate is read by the laser of the CR reader. The effect of the scattering and resultant spreading of the incident laser beam and the emitted luminescence is a source of blurring (0.1-0.2mm) in the image (Fujita et al. 1989).

The observed pinhole PSF is a convolution of the individual blurring functions (Barrett and Swindell 1981):

$$\mathrm{PSF}_{\mathrm{measured}} = \mathrm{PSF}_{\mathrm{focal\ spot}} * \mathrm{PSF}_{\mathrm{pinhole}} * \mathrm{PSF}_{\mathrm{receptor}} \tag{4.1}$$

The presence of $\mathrm{PSF}_{\mathrm{pinhole}}$ indicates that the pinhole itself is a source of blurring in the measured image; we should deconvolve the measured PSF to

*Shavings of bone filed from a lamb roast!

remove its effect. For a pinhole in a clinical test object, whose dimensions are known exactly, it is possible to include the PSF deconvolution as a pre-processing step by modelling the pinhole as a radially symmetric function of constant height and diameter. Provided the size of the pinhole is much smaller than the focal spot[†] then we can reasonably ignore the contribution of pinhole unsharpness.

In the experiments discussed in this chapter the relative contributions vary according to focal spot size and magnification. However, geometric unsharpness is usually the dominant factor, except at low magnification/fine focal spot. Some simple calculations show how the contributions of focal spot and geometric unsharpness contribute to the overall blurring, but for a more thorough treatment the reader should refer to Dance (1988). The magnified focal spot has dimensions $M-1$ as great at the receptor. In the 1.9BF setting the focal spot has dimensions 0.7×0.3mm which become 0.63×0.27mm at the receptor so the focal spot PSF is dominant here. In the 1.9FF setting, with the focal spot dimensions being 0.2×0.2mm the PSF at the receptor is 0.18×0.18mm which is nearer to the size of the contribution of the receptor unsharpness described above. In both the 3.0BF and 3.0FF configurations geometric unsharpness is dominant.

Of course in practice we do not need to know the relative contributions; it is the final composite PSF which generates the observed blurring, and it is this PSF which is measured using the pinhole.

An implicit assumption of the experiments in this chapter are that our test objects are of negligible thickness and that the PSF does not vary across the thickness of the object (i.e. vertically). To minimise this approximation PSFs were obtained (wherever possible) from the centre of the test object thickness.

A note about printed image display

In the following sections we shall make numerous comparisons between images and for display purposes we shall window the images appropriately. Many of the images considered have a uniform background. It is the nature of MEM to drag this background down to very low pixel values. For those MEM images we will window the display of the restored image to an appearance similar to that of the original image. Experience has shown that the visibility of particular features in printed form is better this way. Images which were formally scored by observers (section 4.2.1) were not windowed

[†]The pinhole cannot be so small that x-ray transmission would require the use of a high mAs setting.

in this manner prior to scoring: an appropriate window was chosen for each image by the experienced observers assigning scores.

4.1.1 De-blurring edges and simple objects

The metal phantom was imaged in various configurations on a laboratory-based Siemens Mammomat 3 mammography unit:

1.9BF (Magnification 1.9, broad focus) The plane of the phantom was 33.4cm from the tube focus and 30.6cm from the plane of the receptor. The image would be expected to be unacceptably degraded by geometric blurring.

1.9FF (Magnification 1.9, fine focus) The phantom was in the same position as above, but imaged using the fine focal spot. This is the conventional magnification view provided on this mammography unit and used clinically.

3.2BF (Magnification 3.2, broad focus) The plane of the phantom was 19.8cm from tube focus and 44.2cm from the plane of the receptor. This high magnification factor and focal spot setting is not supported on standard mammography units because the image would be unacceptably blurred.

3.2FF (Magnification 3.2, fine focus) The phantom was in the same position as above, but imaged using the fine focal spot. Again, this high magnification factor is not supported on standard mammography units because the image would be unacceptably blurred.

The radiographic factors used were 50kVp and 40mAs. The image receptor was a Philips ACR-3 computed radiography system with a cassette having dimensions 24cm×30cm. and a pixel size of 152μm (Type B on page 14).

Image Processing

The images obtained from the CR reader were 10-bit SPI format. This obsolete file format is almost the same as the DICOM[‡] standard widely used in medical imaging. Matlab was used to read and process these files directly. Prior to maximum entropy processing the image data, which was in integer format, was scaled by a factor of 2^{10} so that all values lay in the interval [0 1]. For this experiment only small regions, of the image containing objects

[‡]DICOM: Digital Imaging and Communication in Medicine

of interest, were cropped and processed with MEM (Matlab code for this procedure is listed in Appendix D.5). PSFs for the image were obtained from brass pinholes included the plane of each phantom.

The images contained artefacts originating from sources such as dust and dirt on the image receptor to spurious bright pixels at various locations in the images. We will discuss these in more detail in the next section. A non-uniform x-ray illumination also gave the full-size images an obvious varying background brightness. Since we were dealing only with small, cropped regions, we are able to treat the background as approximately constant across each image and make no provision for its effect on the restoration. Again, this will be handled differently when we discuss the processing of full-sized images in later experiments.

Results

Initial maximum entropy deconvolution experiments involved the so-called metal phantom, which was imaged at 1.9BF, 1.9FF, 3.2BF and 3.2FF. The first experiments involved de-blurring of simple edges and small, highly attenuating objects. We will summarise the most important findings of our deconvolution experiments using this phantom.

1. *Simple edges at conventional magnification*

   Figure 4.2 shows part of a section of brass foil imaged in the 1.9BF configuration. The geometric blurring causes a gentle fall-off of pixel intensities between the highly attenuating foil and the empty background. Processing with maximum entropy causes an obvious sharpening of the edge. Intensity profiles across the middle row of each displayed image are shown beneath the images to support the conclusion that processing with MEM has reduced blurring across the edge.[§]

   The blurring of the same edge when imaged in the 1.9FF configuration is shown in figure 4.3. The blurring caused by a fine focal spot is obviously less severe, but again, deconvolution can reverse the some of the degradation created by geometric unsharpness across the edge.

2. *Simple objects at conventional magnification*

   Figure 4.4 a highly attenuating metal washer from the metal phantom imaged under 1.9BF and 1.9FF respectively. The MEM reconstructions are shown alongside the original objects. The reconstructed images show 'ringing' artefacts near sharp intensity transitions (circular

[§]It is tempting to think that 'true' edge of the brass foil should resemble a an ideal 'step-function' but this is probably not the case in reality.

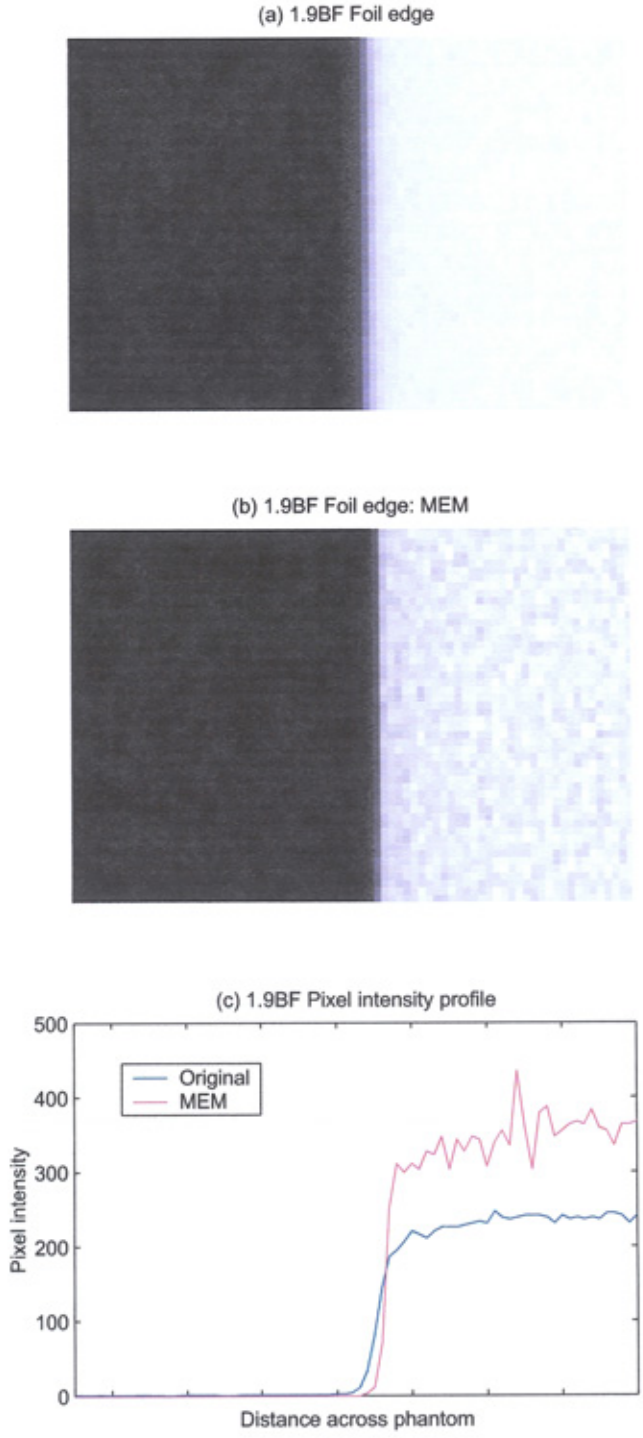

Figure 4.2: Metal phantom: Brass foil edge at 1.9BF. The blurred edge resulting from geometric unsharpness is evident in the original image (a). Processing with maximum entropy, shown in (b), improves the definition of the edge. The pixel intensities across the middle row of each image is shown in (c).

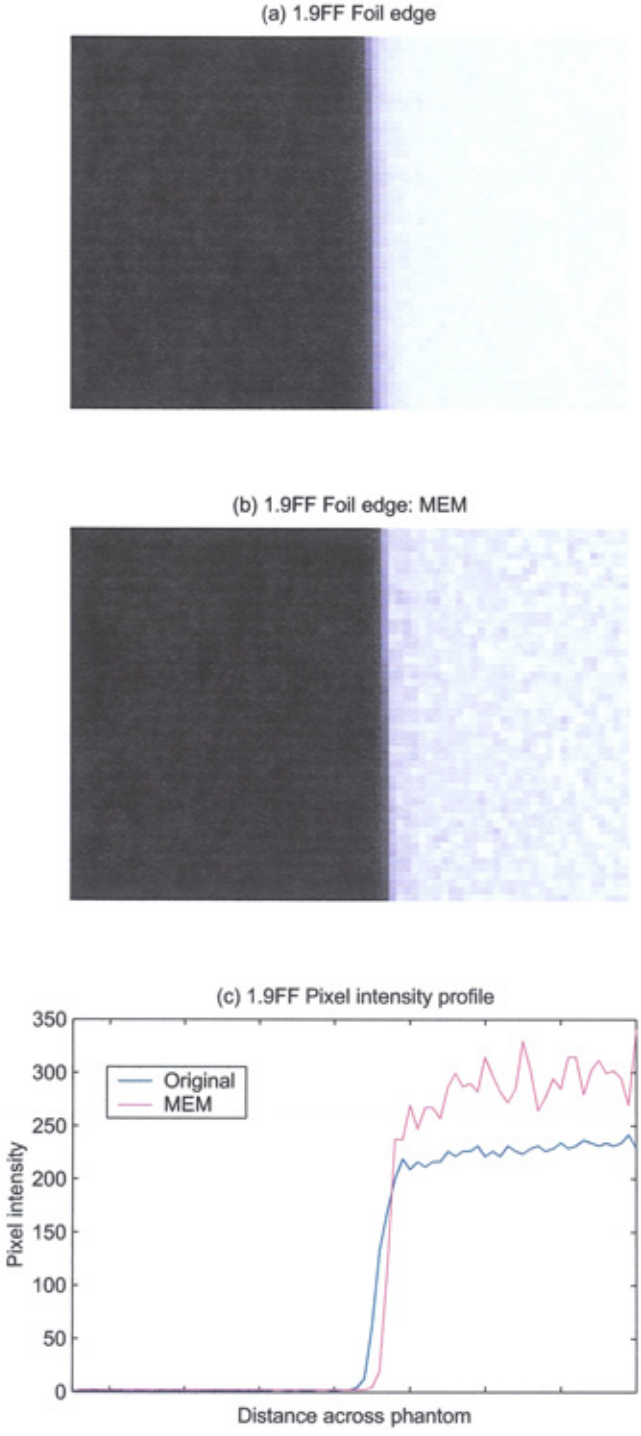

Figure 4.3: Metal phantom: Brass foil edge at 1.9BF. The unsharpness of the edge is evident in the original image (a). Processing with maximum entropy, shown in (b), visibly sharpens the edge. The pixel intensities across the middle row of each image is shown in (c).

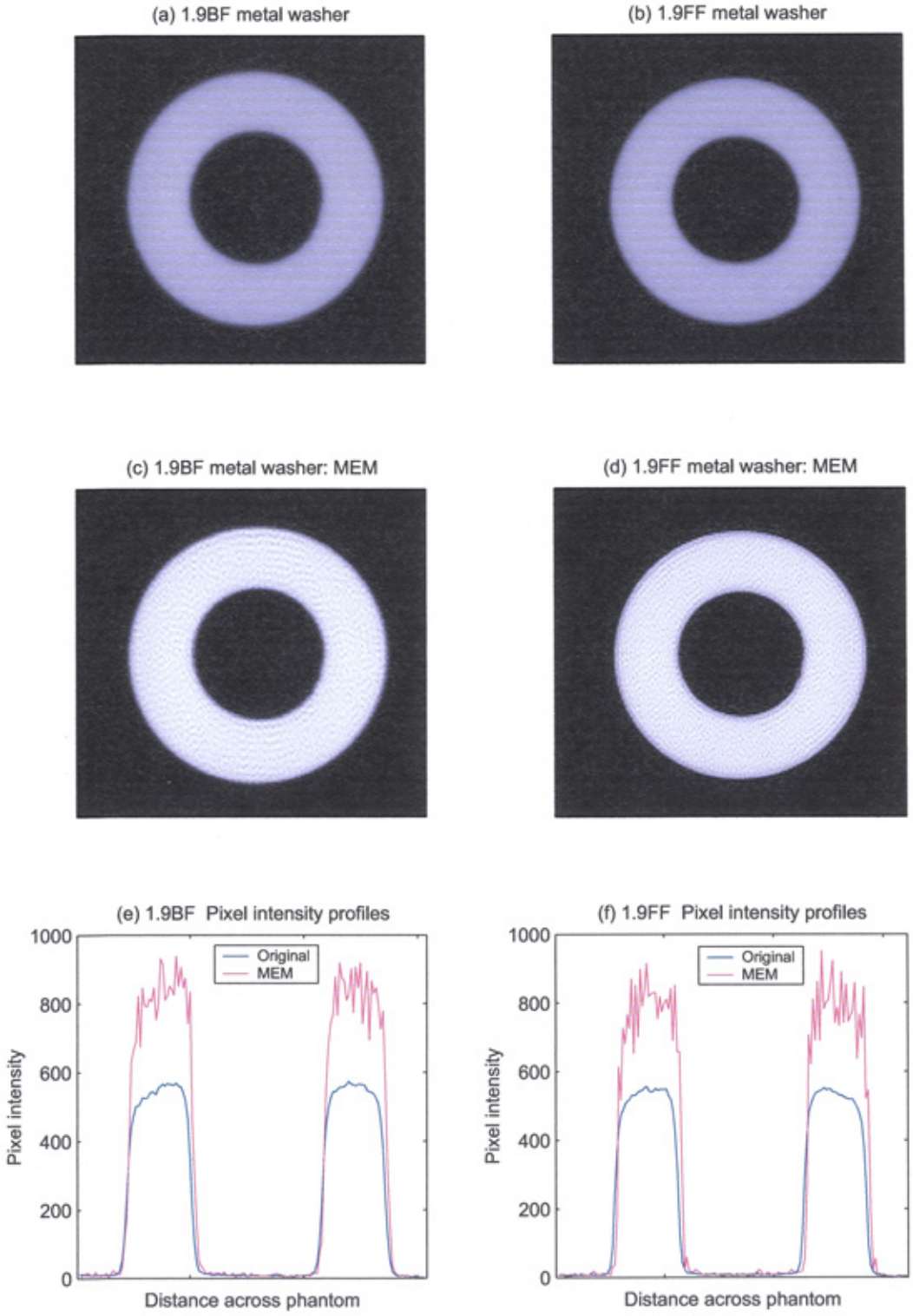

Figure 4.4: A highly attenuating washer shown at 1.9BF and 1.9FF. The focussing effect of deconvolution is present again but the introduction of 'ringing' artefacts near steep intensity transitions are also evident.

patterns on the washer). These artefacts are often seen in regularised deconvolution schemes. The cause of the ringing artefacts is the (usually) poor assumption that the true image distribution is stationary. In a stationary image the statistical content does not vary across the image; comparing the histograms of pixel intensities in different regions will not reveal dramatic differences. In reality most images are non-stationary as some regions will contain edges, and some regions will be light or dark. Lagendijk et al. (1988) address the problem of ringing artefacts in regularised deconvolution schemes and we discuss some of their results, in the context of Historic MEM, here.

The theoretical total error between the true image and the restored image in regularised schemes which assume a stationary image model given by:

$$E_{Tot}(\alpha) = E_1(\alpha) + E_2(\alpha) \tag{4.2}$$

where α is the regularising parameter introduced in section 2.3.5. $E_1(\alpha)$ is a noise magnification error and is a monotonically increasing function of α. $E_2(\alpha)$ is the regularising error and is a monotonically decreasing function of α. The optimal value of α which minimises equation 4.2 leads us to our maximum entropy restoration. The noise magnification error $E_1(\alpha)$ is dependent on global levels of image noise: it acts to globally degrade the restored image. However, the regularising error $E_2(\alpha)$ is a function of the image estimate and is strongly dependent on local structure within the image. Ringing artefacts are associated with this error and should therefore be considered a regularising error.

Ringing can therefore be reduced in a number of ways. In the paper just discussed Lagendijk et al. propose the incorporation of prior knowledge, such as positivity, into the restoration. This can be sufficient in simple astronomical images, for example, where ringing results from the introduction of negative pixel values in the otherwise black background of the restored image. In more complex images, such as those considered in this work, alternative approaches are required as the ringing artefacts are not necessarily caused by negative pixels. Restoration by projection onto convex sets (Youla and Webb 1982; Sezan and Stark 1982) is an alternative approach to deconvolving image data which results in fewer ringing artefacts. A method to reduce ringing artefacts which is applicable to Bayesian methods was discussed by Lantéri et al. (2002).

There is no need to abandon MEM at this stage. While Historic MEM restored images can contain ringing artefacts, the introduction of an ICF into the image model (leading to our Hybrid MEM) can sufficiently reduce these regularising errors.

3. *Simple objects at high magnification*

   The deconvolution of a section containing steel wool imaged under 3.2BF and 3.2FF is shown in figure 4.5. The blurring introduced by the 3.2BF set-up is considerable and although some de-blurring is evident, the restored image cannot satisfactorily recover the resolution lost by geometric unsharpness. Ringing artefacts dominate the restored image. Trial and error experiments showed that stopping the MEM process before the appropriate χ^2 was reached reduced the ringing at the expense of reduced de-blurring.

Discussion

Our original hopes for these early experiments were to test our implementation of maximum entropy and gauge how effectively MEM could reconstruct images of simple objects where the primary cause of image degradation was geometric unsharpness caused by use of a broad or fine focal spot. The results of these initial experiments suggest that MEM can improve the SNR and resolution in most of the situations we investigated.

However, the image blurring associated with higher than standard magnifications, combined with a broad focal spot produced a loss of resolution that could not be recovered using our simplified implementation of MEM. The restored images in the 3.0BF configuration were dominated by restoration artefacts. At this point we ruled out further investigation a broad focus, high magnification imaging set-up.

Although our version of MEM performed very well on the images considered thus far, it was not without problems. Restorations of the brass foil edges showed considerable structure on the surface of the foil (see the pixel profiles in figures 4.2c and 4.3c). The 'jaggedness' of these plots could indicate resolution of the of foil's surface crystal structure. However, it could also be due to an inadequacy of the Historic MEM technique. The use of entropy as a smoothing function does not take into account pixel correlations. Neighbouring pixels can have very different values without affecting the global, maximum entropy of the image. The incorporation of the intrinsic correlation function into MEM is the traditional way to improve this situation (see section 3.4.1).

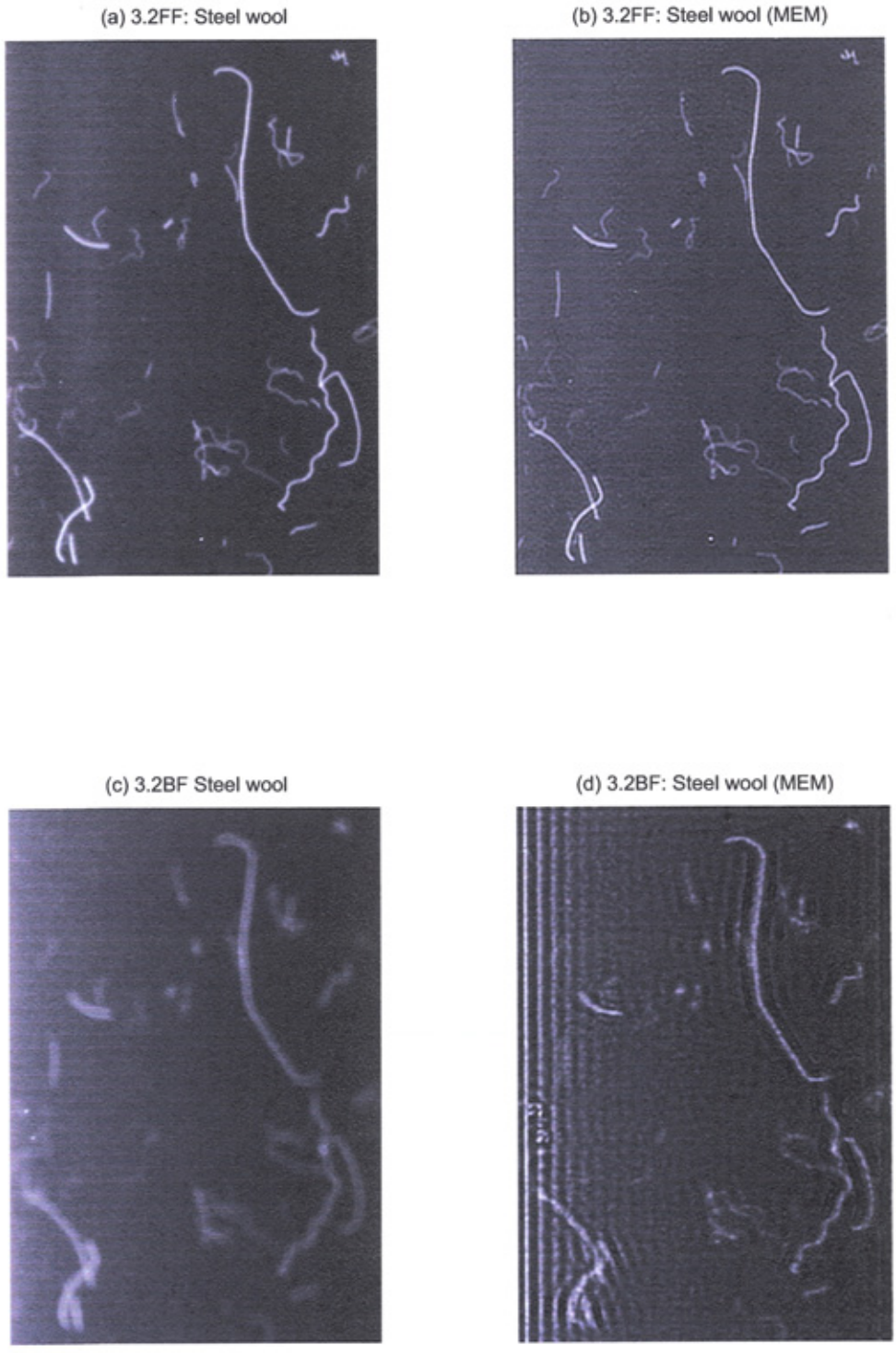

Figure 4.5: Shavings of steel wool imaged at high magnification (3.2x) at fine focus (a) and broad focus (c). The MEM restorations are shown alongside each.

The implementation in Matlab was very intensive in terms of memory requirements and arithmetical operations. For example, the 3.2BF image of steel wool had dimensions of 470 × 325 pixels, or about 3.6% of the total image area from which it was cropped. Despite the relatively small size of this sub-image our Matlab implementation typically required 24 minutes to produce a solution[¶]. Tests on larger images ran correspondingly slower until a point was reached, with images of about 1 million pixels, where the memory requirements were so great that Matlab could not completely process the image.

In order to process larger, or full-sized medical images a better system would be required. The choices we faced were of (a) continuing with our MEM implementation but dividing large images into smaller sub-images for processing, (b) purchasing commercial software and code for use with our images. The problem with (a) was that although we could probably manage to process large images, the quality would probably suffer because discontinuities would be introduced into the image. Also, the time required to process full-sized images was likely to measured in hours rather than minutes. The decision was made at this point to switch to using MEMSYS, a commercial implementation of maximum entropy.

4.1.2 Pre-MEM image processing

The next logical step was to apply maximum entropy deconvolution to test objects which more closely approximated the properties and features of the structures encountered in mammographic imaging. However, some of the early tests with MEM highlighted some potential problems with deconvolving CR images which needed to be addressed before dealing with more realistic images. Before proceeding with descriptions of the next set of experiments, we will discuss some of these problems and the pre-MEM processing stages that enabled better restorations to be obtained. The pre-MEM processing steps were carried out with Matlab.

1. *Removal of a variable background*

   Figure 4.6a shows an image of the bone-fragment phantom obtained under 1.8FF. Images obtained from the MAMMOMAT 3 machine have a background whose brightness varies across the image. The cause of this is the varying intensity of the beam across the film cassette. The variation in beam intensity is a deliberate design consideration

[¶]This involved around 4 minutes to reach a solution for a particular value of α and six iterations to find the appropriate value of α. See figure 3.3

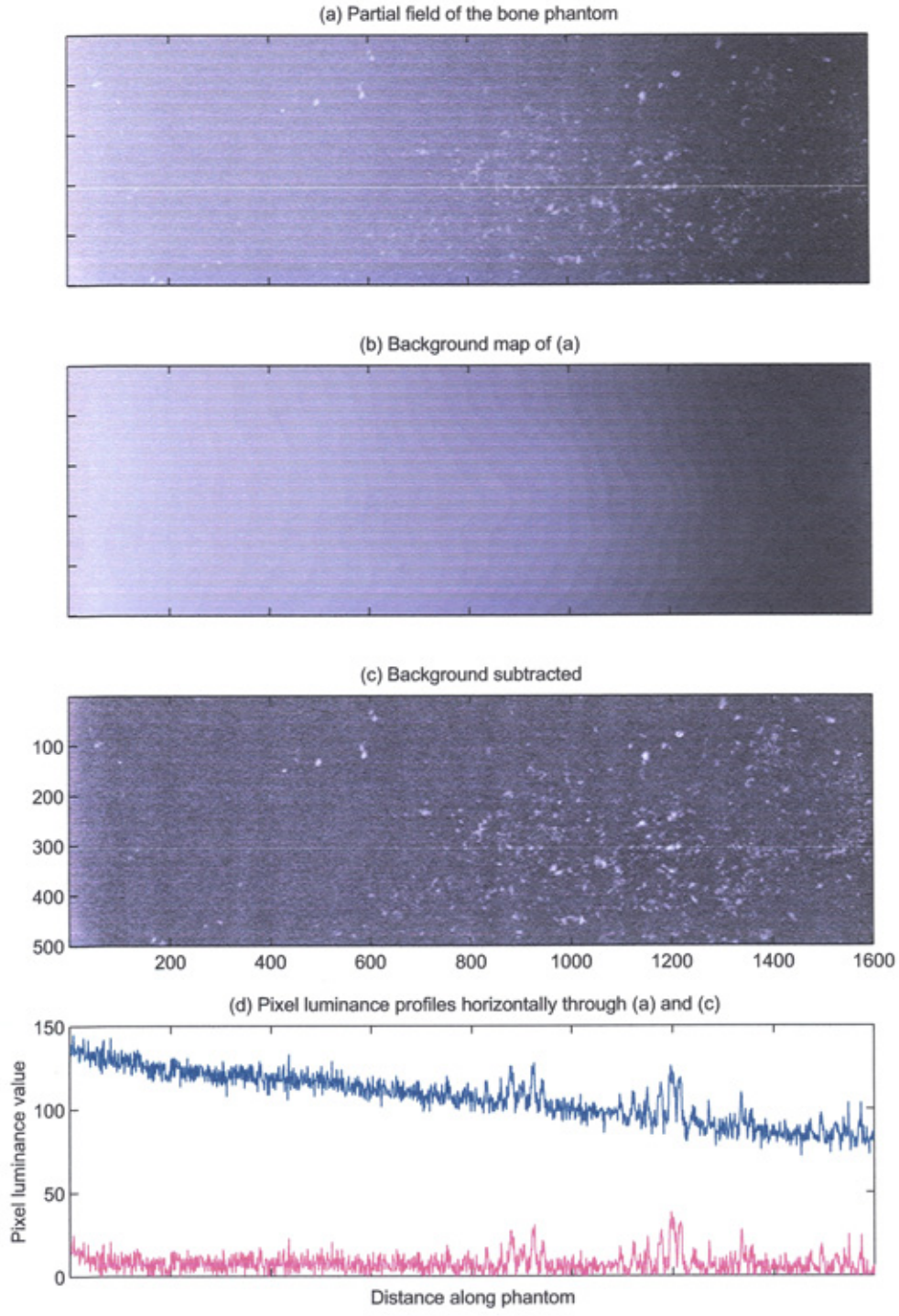

Figure 4.6: (a) part the bone fragment phantom imaged at 1.8FF. The background brightness decreases across the field from left to right. (b) is the low pass background map. (c) is the background corrected image and (d) shows the pixel profiles along the paths indicated in the original (blue) and the background corrected (red).

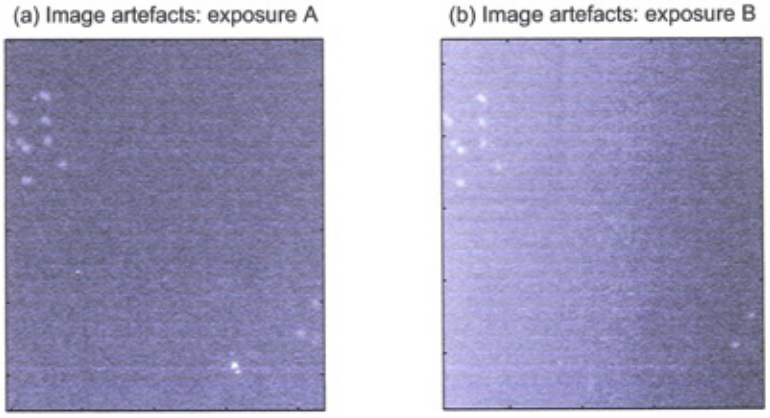

Figure 4.7: Examples of artefacts within images obtained with CR. Both images show the same region of the TORMAM test object. The geometric configuration and exposure settings were the same for each image but a different plate was used to capture the images. The particle groups towards the upper-left and lower-right are real. Deviations from the uniform background elsewhere are artefacts of the CR process.

to compensate for the varying thickness across a real breast tissue. Most of the test objects discussed in this chapter have approximately uniform attenuating properties (on large scales) and so images show a varying brightness in the background. It was desirable for the images to have a uniform background prior to MEM processing (Narayan and Nityananda 1986; Donoho et al. 1992) so a simple scheme to remove the brightness variation ('flat-fielding') is now described.

A Gaussian lowpass filter was constructed and applied to the original images in the Fourier domain to obtain a background map. This map, shown in figure 4.6(b) was subtracted from the original leaving the flattened image of figure 4.6(c), which retained the phantom's high frequency features. The Matlab code given in Appendix D.1 was used for this purpose. We will examine the effect of this procedure on the nonlinear relationship between pixel values and radiation exposure[‖] in section 3.

2. *Removal of spurious pixels*

   Many of the images obtained by the imaging process show artefacts

[‖]See page 14

which have no relation to the objects being imaged: they are a by product of the image acquisition process. In some circumstances the presence of image artefacts in CR can lead to confusion with features such as microcalcifications. A pictorial guide to CR artefacts and their causes was presented by Cesar et al. (2001). In figure 4.7 we show some of these artefacts found in our CR images. The same region of a test object was imaged under the same geometric configuration and exposure settings. The CR plate was different. In image 4.7a there are a number of spurious 'bright' pixels whose values are far above the background. At the same position in image 4.7b there are no such features. A more subtle example of an artefact are the vertical streaks down the centre of image 4.7b. Again, these do not correspond to real features and are absent from image 4.7a.

Artefacts are typically caused by the presence of dust or dirt in the image receptor. Inspection of many artefacts show that there is no characteristic blurring associated them - they have not been degraded by the system PSF.

Removal of CR artefacts can be addressed by incorporating their presence into the model formulation (Highnam et al. 1999). In this work we will not attempt to remove all the artefacts within each CR image. However, the most straightforward of the artefacts to remove are those spurious pixels seen in figure 4.7b. Attempting to deconvolve an image containing isolated bright pixels results in a restoration with strong ringing around sites of the original artefacts. Simple enhancement techniques such as median filtering would remove a lot of these spurious pixels but at the cost of reduced resolution and possible loss of important detail at the smallest scales of the image. To avoid complication of the forward map we simply try to remove such artefacts manually, through a scheme of identifying obvious outliers in the histogram of the region of interest and then resetting the offending pixel intensities to a mean background level. It is also worth pointing out that the plates used to obtain many of the images in this chapter are old and well-used and would no longer be suitable for clinical use. The number of artefacts is rather higher than might be expected.

3. *Linearisation of the data*

   The images obtained from the CR reader were in Siemens SPI file format: 12-bit grayscale images holding 10-bits (1024 gray levels) of image data. As has been noted this format resembles closely a standard image format called DICOM and Matlab was able to read the image

data directly. The images were linearly mapped into the range [0 1] prior to further processing.

It was stated in section 1.3.6 the pixel values in the image are related non-linearly to the number of x-ray photons:

$$I_{\log} = \left(\frac{1024}{L}\right) \log I_{\text{lin}} - \log\left(\frac{s}{200}\right) + 511 \tag{4.3}$$

One of the assumptions which underpin the validity of our maximum entropy arguments of Chapter 3 is that the pixel values are linearly representative of photon counts E impinging the detector at the position of the image pixel. Images obtained from the CR reader should therefore be linearised prior to further processing. We use the following equation to do this:

$$I_{\text{lin}} \propto 10^{I_{\log} L/1024} \tag{4.4}$$

The images obtained in our mammography work typically have $L \approx 1.6$.

Many of the images under consideration in the following experiments are simple, being comprised of the object signals embedded in a relatively high 'background' signal, whose mean level is not much lower than the signal. With a slight change of notation we attach superscripts to the exposure I_{lin} and define relationships between pixels containing object signal and background signal. As we wish to discriminate between the contributions of the object and the background we let $I_{\text{lin}}^{\text{image}} = I_{\text{lin}}^{\text{back}} + I_{\text{lin}}^{\text{obj}}$ and write:

$$\begin{aligned} I_{\log}^{\text{image}} &= \left(\frac{1024}{L}\right) \log I_{\text{lin}}^{\text{image}} - \log\left(\frac{s}{200}\right) + 511 \\ &= \left(\frac{1024}{L}\right) \log(I_{\text{lin}}^{\text{back}} + I_{\text{lin}}^{\text{obj}}) - \log\left(\frac{s}{200}\right) + 511 \end{aligned} \tag{4.5}$$

The pixel values of the background signal are similarly given by:

$$I_{\log}^{\text{back}} = \left(\frac{1024}{L}\right) \log I_{\text{lin}}^{\text{back}} - \log\left(\frac{s}{200}\right) + 511 \tag{4.6}$$

Our pre-processing procedure was used to construct a background map and subtract it from the original image. We will investigate the effect

of this on the validity of our MEM solutions. Background subtraction yields:

$$\begin{aligned} I_{\log}^{\text{image}} - I_{\log}^{\text{back}} &= \left(\frac{1024}{L}\right)\left[\log(I_{\text{lin}}^{\text{back}} + I_{\text{lin}}^{\text{obj}}) - \log I_{\text{lin}}^{\text{back}}\right] \\ &= \left(\frac{1024}{L}\right)\left[\log\left(\frac{I_{\text{lin}}^{\text{back}} + I_{\text{lin}}^{\text{obj}}}{I_{\text{lin}}^{\text{back}}}\right)\right] \\ &= \left(\frac{1024}{L}\right)\left[\log\left(1 + \frac{I_{\text{lin}}^{\text{obj}}}{I_{\text{lin}}^{\text{back}}}\right)\right] \end{aligned} \tag{4.7}$$

Now the ratio of object signal to background signal is of the order of 0.1-0.2 in many cases (see figure 4.17, for example). In these circumstances we will argue that the data can be effectively linearised by the simple background subtraction process seen earlier. We know for small x that $\ln(1+x) \approx x$. Similarly, through a change of logarithmic base we see that $\log(1+x) = \ln(1+x)\log e$. So for small x we have the relationship $\log(1+x) \sim x$. We invoke this approximation in equation 4.7 to obtain:

$$I_{\log}^{\text{image}} - I_{\log}^{\text{back}} \sim \frac{I_{\text{lin}}^{\text{obj}}}{I_{\text{lin}}^{\text{back}}} \tag{4.8}$$

Therefore, under certain circumstances the background removal effectively linearizes the relationship between image pixel values and the numbers of x-ray photons striking the image receptor. The maximum entropy analysis of chapter 3 requires the data to be linearly related to the photon counts for the forward map to be valid.

4.1.3 Tests with Bone and Eggshell phantoms

The next series of experiments involved the use of test objects whose physical properties were closer to those of the biological features in a real breast. For this, we constructed the Bone Phantom and the Eggshell phantom, described at the start of section 4.1. With these phantoms we obtained images using radiographic settings (operating voltage, tube current) approaching those used in real mammographic imaging.

Experimental method

The Bone and Eggshell phantoms were imaged in various configurations on a laboratory-based Siemens Mammomat 3 mammography unit. The focal spot to image receptor distance was measured as 64cm.

2.0BF (Magnification 2.0, broad focus) The plane of the phantom was 32cm from the tube focus and 32cm from the plane of the receptor. The image would be expected to be unacceptably degraded by geometric blurring.

2.0FF (Magnification 2.0, fine focus) The phantom was in the same position as above, but imaged using the fine focal spot. This is the almost the conventional magnification view used clinically.

4.0FF (Magnification 4.0, fine focus) The plane of the phantom was 16cm from tube focus and 48cm from the plane of the receptor. This high magnification factor and focal spot setting is not supported on standard mammography units because the image would be unacceptably blurred.

The radiographic factors used to obtain the images were 28kVp, 16mAs. The image receptor was a Philips ACR-3 computed radiography system utilising an image cassette with dimensions 18cm×24cm with a pixel size of 101μm (Type C on page 14).

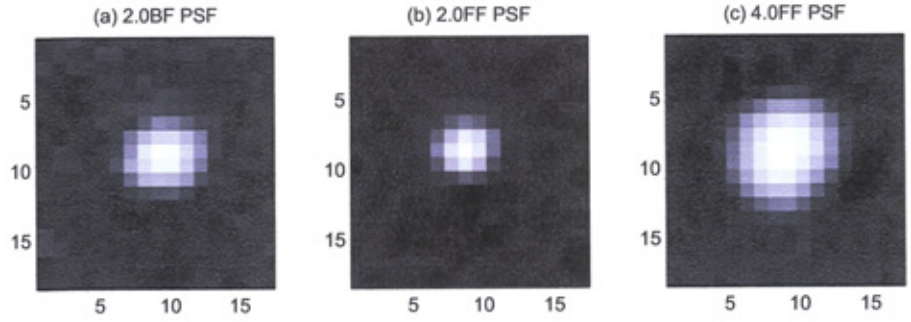

Figure 4.8: Pinhole images of the PSF obtained in the three imaging configurations. (a) is a broad focal spot obtained at magnification 2.0. It is noticeably larger than the fine focal spot shown in (b). Although the focal spot shown in (c) is also fine focus, it has been geometrically magnified and is greater in extent than even the broad focal spot PSF in (a).

In order to provide a sample of the point spread function (PSF) in the plane of the phantom, a brass foil with a pinhole was included in the test images. Including the PSF with the image was convenient in this demonstration as the scaling and interpolation step was avoided. Examples of raw PSF images are shown in figure 4.8.

Image processing procedure

Pre-MEM processing was carried out in Matlab. The non-uniform background was subtracted from the original image and spurious bright pixels were also removed. Images were saved as 16-bit TIFF to prevent loss of information and to ensure compatibility with MEMSYS. The images were treated with the Hybrid MEM scheme described in section 3.4. This was achieved in MEMSYS by using the Historic Maximum Entropy switch and defining an appropriate Intrinsic Correlation Function. The default image model used for these experiments was a uniform image with low values ($m_i = 0.001$ everywhere); this was to encode our belief that, in the absence of data, the assumption is that no x-ray absorption has taken place. The restored images generated by MEMSYS were formatted as 8-bit PNG.

Results

Figures 4.9 and 4.10 show the same region of the bone phantom (measuring about 8.8mm×8.8mm). These displayed images were cropped from the much larger processed images; we show just this small section so that the geometric blurring around the numerous bone fragments is visible. To demonstrate some of the effects of applying MEM we also present some pixel intensity profiles through some of the bone and eggshell fragments in the original and restored images. The original images are have a grainy appearance due to random fluctuations in noise. It is seen that the restored images are much smoother in appearance and is well shown in the pixel profiles. The strength of the peaks has been increased by MEM processing whereas the background (and noise) has been moved towards the default image model (uniform, low values). The intensity profile in these examples are taken along a row passing through one or more fragments. Estimates of the SNR were obtained from a bone fragment in each image. For the 2.0BF configuration we found $\mathrm{SNR}_{\mathrm{orig}} = 5.1$ and $\mathrm{SNR}_{\mathrm{MEM}} = 35.7$. For the same fragment in the 2.0FF configuration we found $\mathrm{SNR}_{\mathrm{orig}} = 6.3$ and $\mathrm{SNR}_{\mathrm{MEM}} = 41.3$

Figures 4.11 and 4.12 show the same region of the eggshell phantom (measuring 8.8mm×8.8mm). The displayed images were cropped from the much larger processed images; we show just this small section so that the

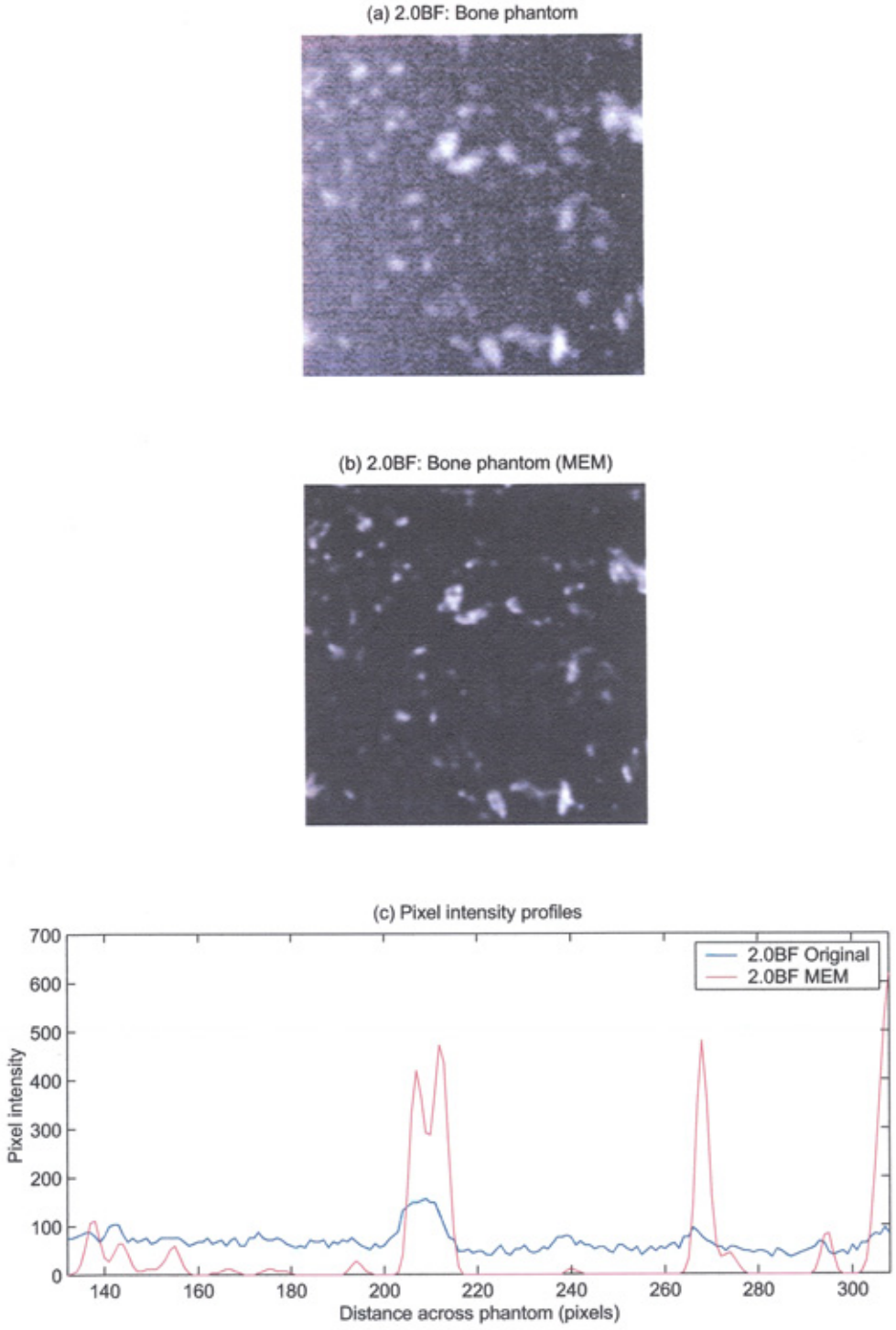

Figure 4.9: 2.0BF: Comparison of (a) the original and (b) MEM processed images of a section of the Bone phantom (c) Pixel intensity profile through several of the fragments.

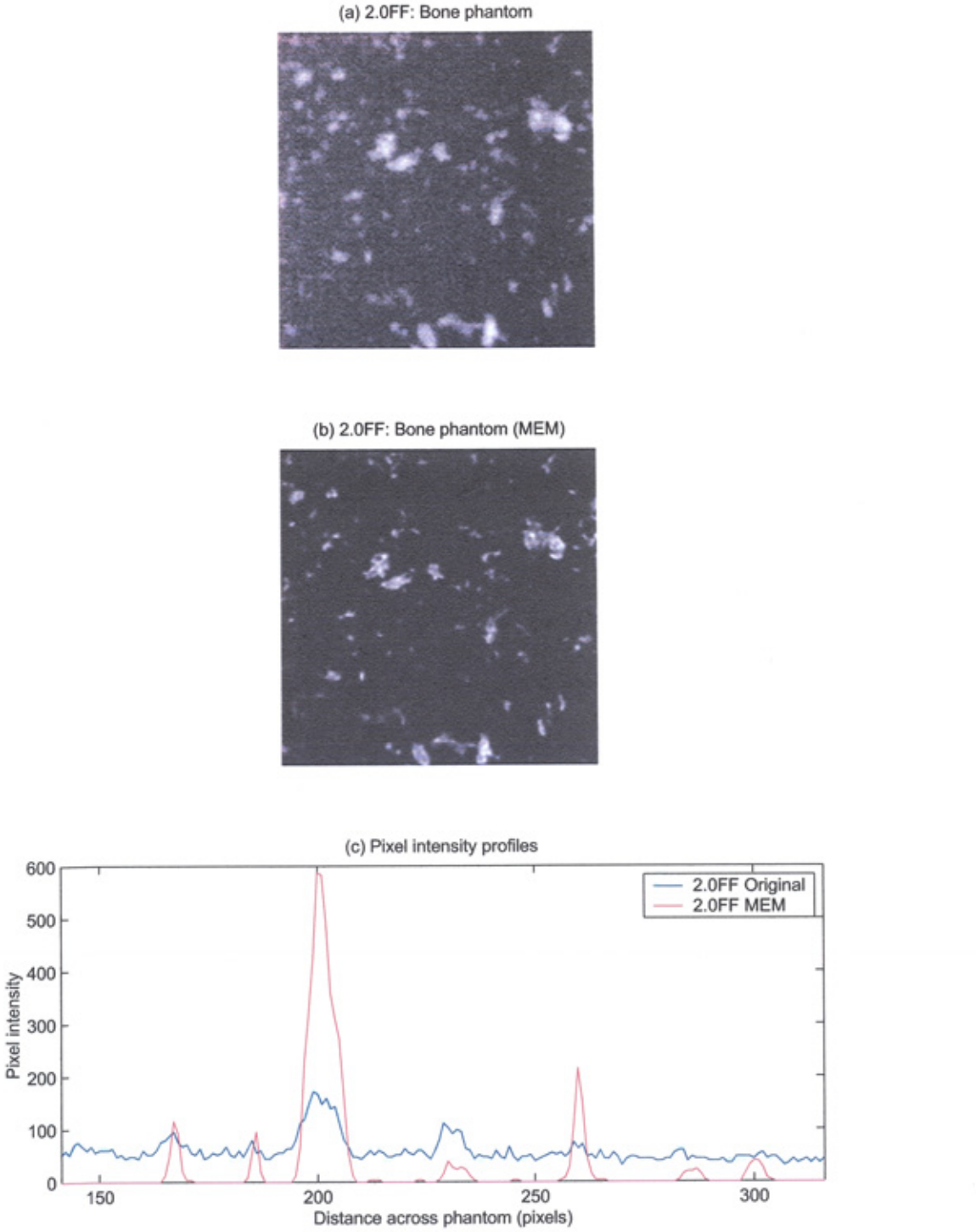

Figure 4.10: 2.0FF: Comparison of (a) the original and (b) MEM processed images of a section of the Bone phantom (c) Pixel intensity profile through several of the fragments.

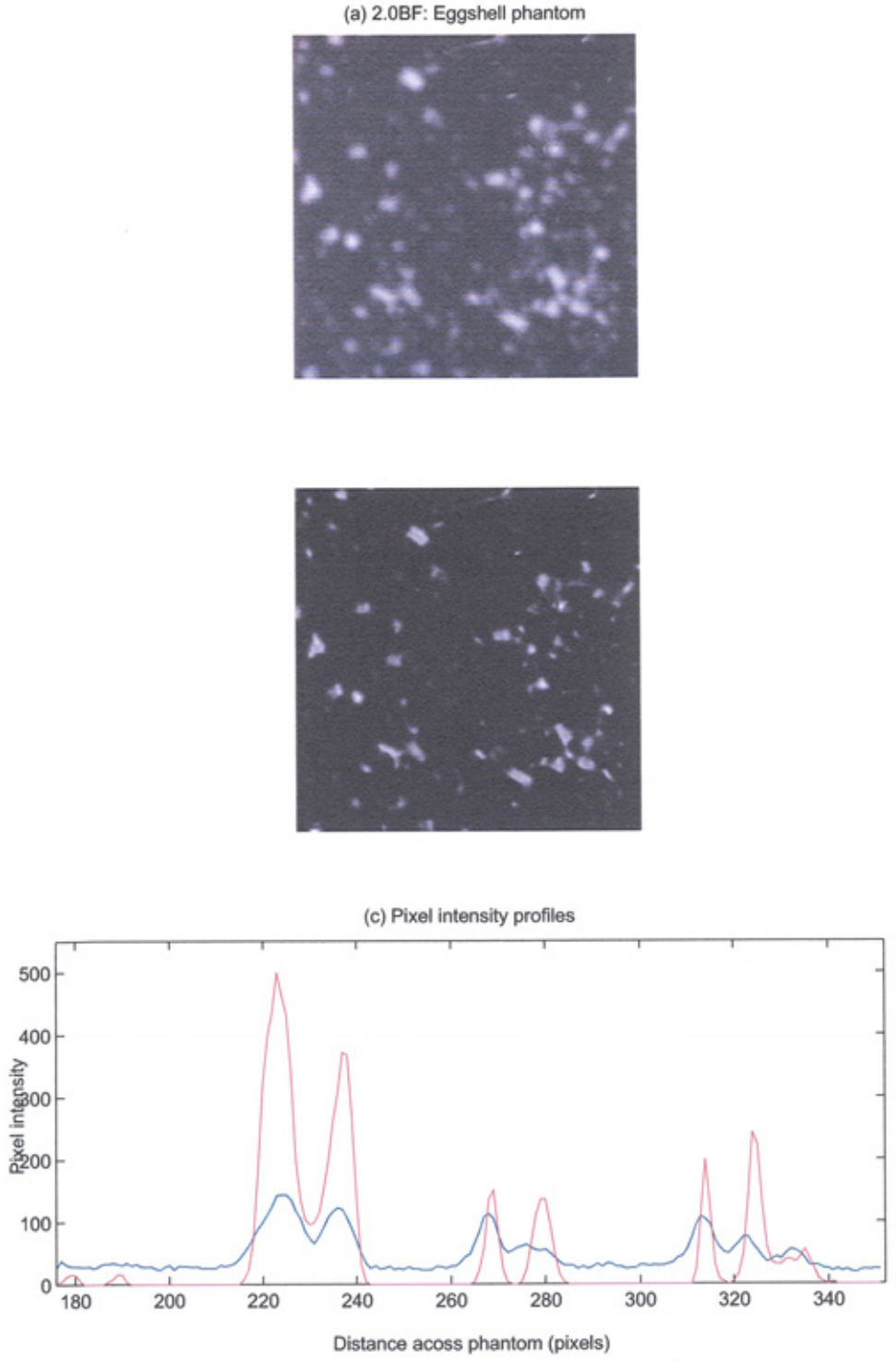

Figure 4.11: 2.0BF: Comparison of (a) the original and (b) MEM processed images of a section of the Eggshell phantom.

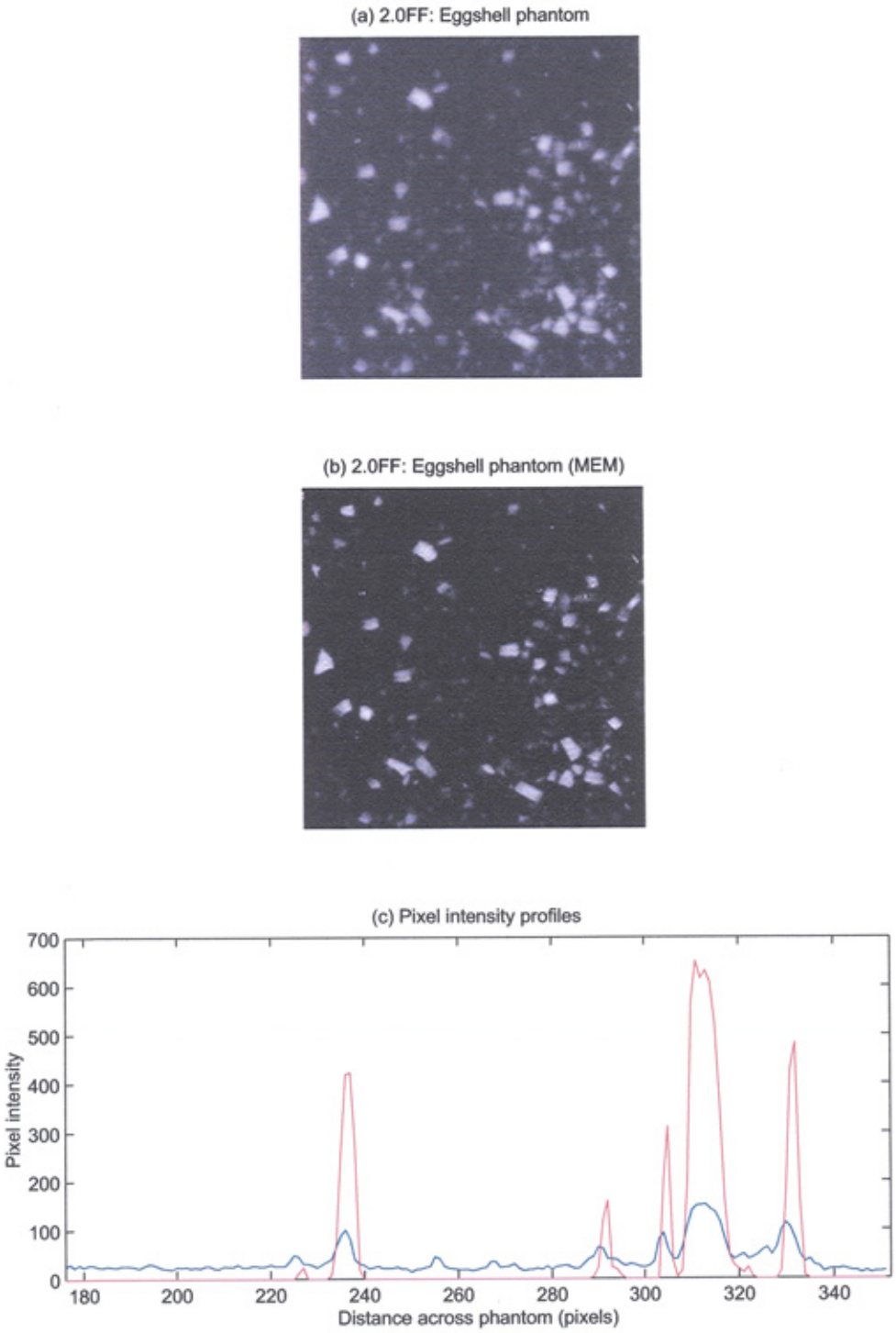

Figure 4.12: 2.0FF: Comparison of (a) the original and (b) MEM processed images of a section of the Eggshell phantom.

geometric blurring is visible. As with the bone phantom images, also included in each figure is a profile of the pixel intensity along a row passing through one or more eggshell fragments. Estimates of the SNR were obtained from the same fragment in each image. For the 2.0BF configuration we found $SNR_{orig} = 8.5$ and $SNR_{MEM} = 55.9$. For the same fragment in the 2.0FF configuration we found $SNR_{orig} = 9.3$ and $SNR_{MEM} = 56.0$

We proceed now to images obtained with the 4.0FF configuration. Figure 4.13 shows a small region (4.0mm×3.7mm) of the bone phantom and figure 4.14 shows a small section (7.8mm×6.0mm) of the eggshell phantom. The higher magnification offers a much greater spatial resolution of the object

Discussion

Images obtained in this experiment showed an obvious reduction in geometric blurring. The overall appearance of the restored images is that they have been focussed; edges appear sharper and the signal content of the images is boosted 'higher' above the background and noise. Examination of various pixel profiles through features in the images show this to be the case.

The images obtained with 2.0BF and 4.0FF show the most dramatic change in terms of blur reduction. This is to be expected as the PSF size for these set-ups is significantly larger than for the 2.0FF configuration. Some improvement in the sharpness of features is evident in the 2.0FF images but most of the improvement is due to MEM dragging the background towards the default model and smoothing the noise across the image.

The estimates of SNR taken alongside the observed quality of the MEM restorations point to another advantage of applying MEM deconvolution to this imaging situation. The 2.0BF images rivalled or surpassed the quality of the unprocessed 2.0FF. The relative merits of a broad focal spot over a fine focal spot were discussed in the introductory chapter (see page 26).

In this experiment the PSF was obtained from a brass foil pinhole in the same plane as the object, i.e. a pinhole projection of the focal spot. In any practical application of this technique, the PSF information would be obtained from calibration images appropriately scaled for the position of the object. Including the PSF with the object was convenient in this demonstration as the scaling and interpolation step was avoided.

4.2 TORMAM phantom experiments

The Bone and Eggshell phantom experiments showed that MEM could reduce blurring of images of objects whose size and composition were approaching

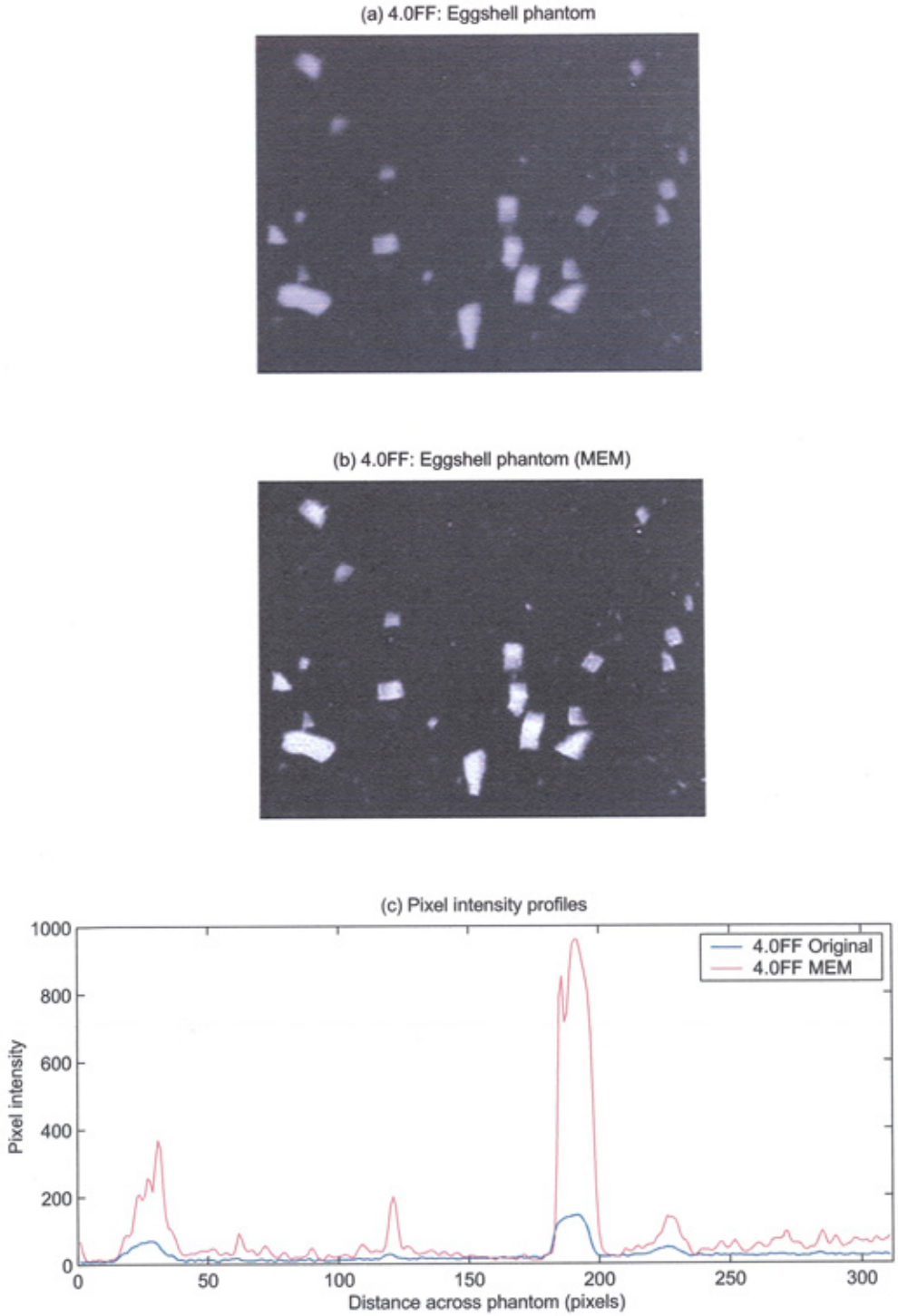

Figure 4.14: 4.0FF: Comparison of (a) the original and (b) MEM processed images of a section of the Eggshell phantom.

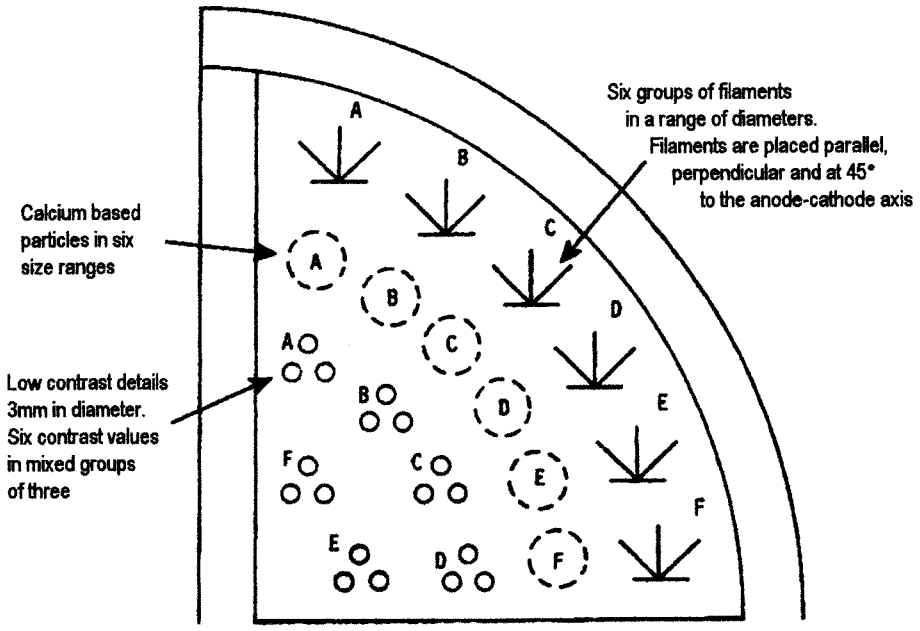

Figure 4.15: Layout of the quantitative side of the Leeds TORMAM test object

the properties of breast tissue and associated structures. The next stage in the process was to use a clinically recognised test object to gauge, in a more quantitative way, how effectively MEM was improving the images.

Image quality comparisons were carried out using the Leeds TORMAM test object (figure 4.15) at various settings of geometric magnification. This phantom contains three groups of test features; fibres, simulated microcalcification clusters and low contrast plastic disks, plus an area designed to give an anthropomorphic impression of a breast parenchymal pattern with overlying microcalcification clusters (Cowen et al. 1992). This latter area of the phantom was not used in these experiments.

The TORMAM phantom is 1.1cm thick and normally placed on top of a stack of D-shaped Perspex plates 3.5cm thick. When imaged at 28kVp this test object, with a total thickness of 4.6cm is approximately equivalent to breast of thickness 5.3cm with a glandularity of 29% (Dance et al. 2000a).

4.2.1 Reduction of blurring experiments**

Experimental method

The Leeds TORMAM phantom was imaged in various configurations on a laboratory-based Siemens Mammomat 3 mammography unit, with measured focal spot sizes of 0.7×0.3mm broad focus and 0.2×0.2mm fine focus. In all cases the phantom was imaged on top of a stack of Perspex to provide realistic scatter and attenuation as recommended in the instructions for the phantom. The radiographic factors used were 28kVp, 40mAs. These factors were chosen to be representative of the values used in routine mammographic quality assurance tests. The level of quantum noise in the test images was therefore realistic. The image receptor was a Philips ACR-3 computed radiography system with a cassette having dimensions 18cm×24cm and a pixel size of 101μm (Type C on page 14).

Three imaging geometries were used to illustrate varying degrees of focal spot geometrical blurring:

1.8BF (Magnification 1.8, broad focus) The plane of the phantom was 33.5cm from the tube focus and 27cm from the plane of the receptor. The image would be expected to be unacceptably degraded by geometric blurring.

1.8FF (Magnification 1.8, fine focus) The phantom was in the same position as above, but imaged using the fine focal spot. This is the conventional magnification view provided on this mammography unit and used clinically.

3.0FF (Magnification 3.0, fine focus) The plane of the phantom was 20.5cm from tube focus and 40cm from the plane of the receptor. This high magnification factor is not supported on standard mammography units because the image would be unacceptably blurred.

A piece of brass with a small pinhole was positioned alongside the TORMAM phantom and was used to obtain an image of the focal spot (see page 108). The raw PSFs obtained with each configuration are shown in figure 4.16. In the above list only the 1.8FF set-up is used in practice; the amount of geometric blurring introduced in an image is too great with the other configurations. The intention in this section is to explore the possibility of also using the unconventional set-ups 1.8BF and 3.0FF in conjunction with MEM to reduce the associated blurring.

**A summary of this work has been published (Jannetta et al. 2004)

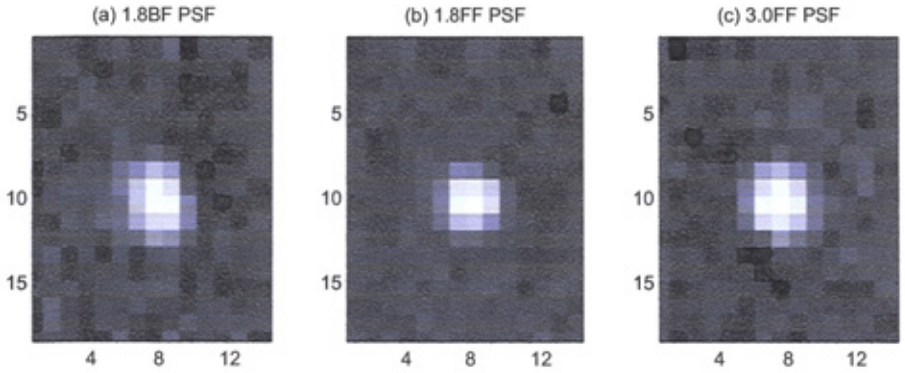

Figure 4.16: Pinhole images of the PSF obtained in the three imaging configurations. 1.8BF PSF (a) is noticeably broader than the 1.8FF PSF (b). Although the 3.0FF PSF is also fine focus it has been geometrically magnified and is comparable in size to (a).

The original and processed sets of images were viewed and scored by two independent observers both experienced in the use of mammographic image quality test phantoms. The test images were graded using the 3, 2, 1, 0 scoring system recommended in the TORMAM phantom instructions and adopted in surveys of mammographic image quality in the UK Breast Screening Programme (Young and Ramsdale 2003).

Image processing procedure

The images obtained from the CR reader were in Siemens SPI file format which were 12-bit grayscale images holding 10-bit image data. Matlab R12 was used to read these images and perform some preprocessing so that the subsequent MEM processing could be applied. These steps were the same as in the previous experiment (see page 96)

- Subtraction of a background map to flat-field the image.
- Removal of spurious, bright pixels.
- Cleaning the PSF images; in this case the noisy background of the PSF was filtered out (typically by discarding those pixels with less than 10% of the PSF peak value). This had the effect of slightly narrowing each PSF - thus leading to a conservative under-restoration.

MEMSYS5 was used to treat the images following the Matlab preprocessing. The interface to the MEMSYS5 kernel accepts image and PSF-image

files as inputs (Matlab image format was convenient) and allows the setting of certain parameters related to the theory described in Chapter 3. The default image model m was defined to be a flat image with low pixel luminance values ($m_i = 0.001$ for all i). This codifies our belief that, in the absence of data, very high photon counts were recorded (i.e. no absorption due to intervening material). The program's Historic MEM switch and setting an ICF (of width 1) was done to implement the Hybrid MEM scheme of section 3.4.

The processed images contained between 2.2 million and 3.9 million pixels. MEMSYS5 typically converged to a solution within 15 α-iterations with a processing time of four to eight minutes, for an image of the full test object shown in figure 4.15. The smaller images to be presented in figures 4.18 to 4.23 were cropped from such an image after processing, rather than processed individually. The processed output files from MEMSYS5 were 8-bit PNG files. All image processing was carried out on a Pentium 4 2.4GHz machine with 512MB of RAM.

Results

Before proceeding to a systematic comparison of the original and MEM processed images and related scores, we will say a few words about our initial aspirations for this technique. As mentioned at the beginning of this chapter (see page 84) these were to show improved spatial resolution without reduction in signal-to-noise ratio, particularly a level of performance in which MEM processed 1.8BF images are at least as good as unprocessed 1.8FF ones. Our belief is that such a performance would be of clinical interest.

Figures 4.18 and 4.19 compare appropriate images of crow's feet and simulated microcalcifications. These figures show that MEM reconstruction can give improvements in both resolution *and* perceived signal-to-noise ratio, and that the MEM processed 1.8BF images are marginally better in this instance than the 1.8FF unprocessed ones in both respects. This promise encouraged us to undertake a systematic evaluation, and to consider magnifications greater than those normally used in clinical practice. Three particle groups (B,D and E) were chosen for this evaluation. The nominal particle size ranges for these groups are 180-283 μm, 106-177 μm and 90-141 μm respectively (Cowen et al. 1992). In addition to the scoring system mentioned earlier we also calculate some signal to noise ratios (SNR) for these particular features.

Figure 4.20 shows images of the three particle groups, cropped from the images taken under 1.8BF conditions. The selected particle group is clearly visible in figures 4.20a and 4.20d, discernible in figure 4.20e, just discernible

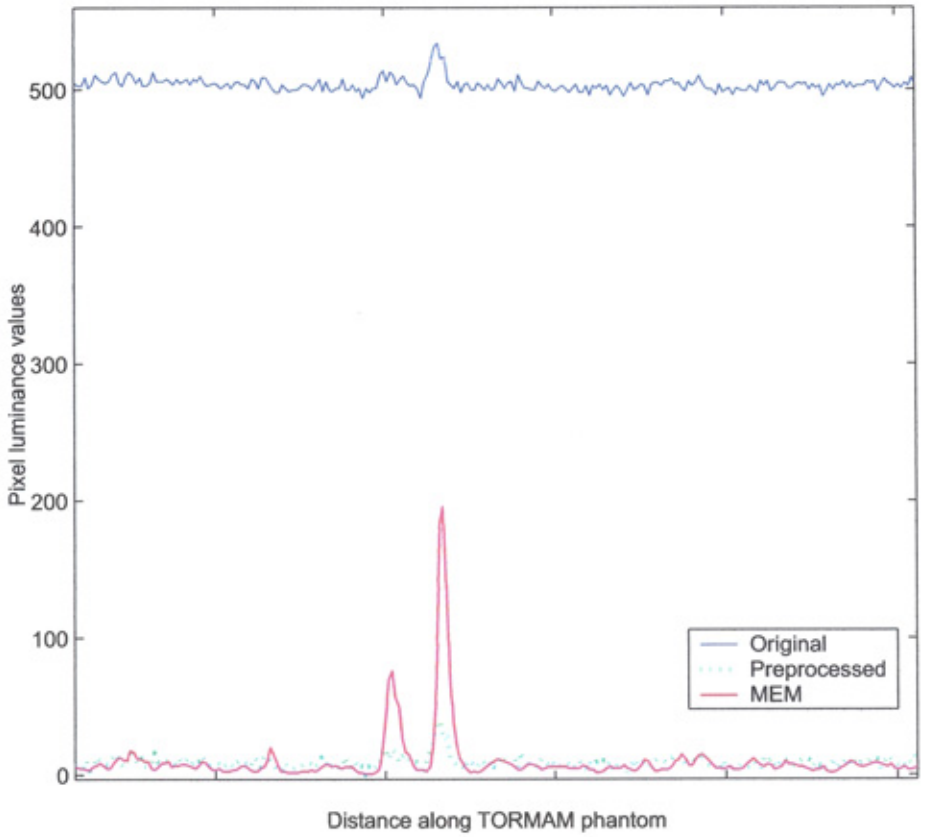

Figure 4.17: 1.8BF. Pixel intensity profiles through particles in group A. The preprocessing step reduces and flattens the varying background of the image but leaves the structure and noise intact. MEM processing smoothes the noise and increases the signal-to-noise ratio of features deemed statistically to be caused by real objects.

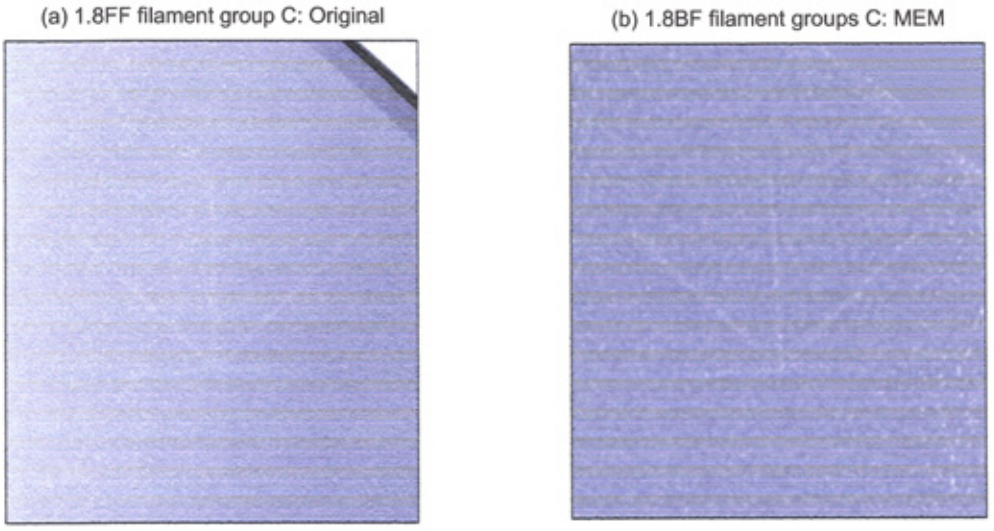

Figure 4.18: Comparison of the filament group C imaged under 1.8FF with the same feature imaged under 1.8BF and processed with MEM.

in figure 4.20b and not seen in figures 4.20c and 4.20f; however, we show the latter images to preserve the two by three format, which eases comparison with later improvements. As expected, unprocessed images obtained in this configuration are too blurred to be clinically useful. However, the MEM processed images show significant improvements in resolution and signal-to-noise ratio, particularly regarding the high frequency noise, which were quantified using pixel profiles similar to that shown in figure 4.17; the quoted SNR is the difference between the largest signal and the mean background within each group, divided by the standard deviation of the signal in a region close to the group. Figures for Group B are $SNR_{orig} = 9.0$ and $SNR_{MEM} = 50.3$. Group D: $SNR_{orig} = 5.9$ and $SNR_{MEM} = 13.7$. Particle group E is not detected at this setting.

Figure 4.21 shows cropped images of the same three particle groups, obtained with the conventional 1.8FF configuration. Remarks regarding group visibility are as for figure 4.20. The pinhole PSF imaged under these conditions is small and approximately Gaussian in shape, comprising just a few pixels. As expected, images obtained in this configuration are sharper than in the 1.8BF case; nevertheless MEM restoration still shows significant improvements. SNR measurements for group B are $SNR_{orig} = 9.0$ and $SNR_{MEM} = 49.8$. Group D: $SNR_{orig} = 5.4$ and $SNR_{MEM} = 18.5$. Particle group E is not detected at this setting.

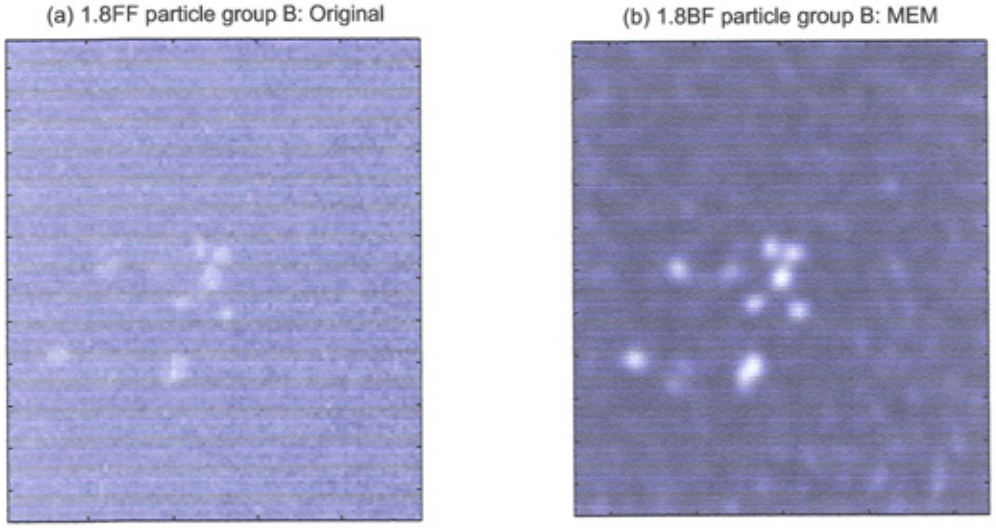

Figure 4.19: Comparison of particle group B imaged under 1.8FF with the same feature imaged under 1.8BF and processed with MEM.

Figure 4.22 shows cropped images of the same three particle groups, obtained with an unconventional 3.0FF configuration. In this case all three groups are detectable in the original image of the phantom, but the MEM processed images show clear improvements in resolution, enabling fine details of individual microcalcifications to be discerned in image 4.22d. SNR measurements for group B are $SNR_{orig} = 8.8$ and $SNR_{MEM} = 49.8$. Group D: $SNR_{orig} = 5.8$ and $SNR_{MEM} = 22.4$. Group E: $SNR_{orig} = 5.2$ and $SNR_{MEM} = 10.4$.

To effect the same comparison as that illustrated in figure 4.19, between unprocessed 1.8FF images and MEM processed 1.8BF ones, the bottom row in figure 4.20 should be compared with the top row in figure 4.21.

Figure 4.23 shows images of filament groups B, D and F taken with the 1.8FF configuration. These respectively have diameters 0.35, 0.25 and 0.20 mm, length 10 mm. All three groups are detectable in the original and MEM processed images, but again the latter show improvements in resolution and visibility. To avoid information overload we do not show the 1.8BF and 3.00FF cases here, but experience with these matches that with the particle groups, and is quantified in the table of scores discussed below.

The results obtained from scoring the original and MEM processed images under each of the three imaging geometries are shown in Table 4.1. The processed images show an improved score in each case. This implies that the

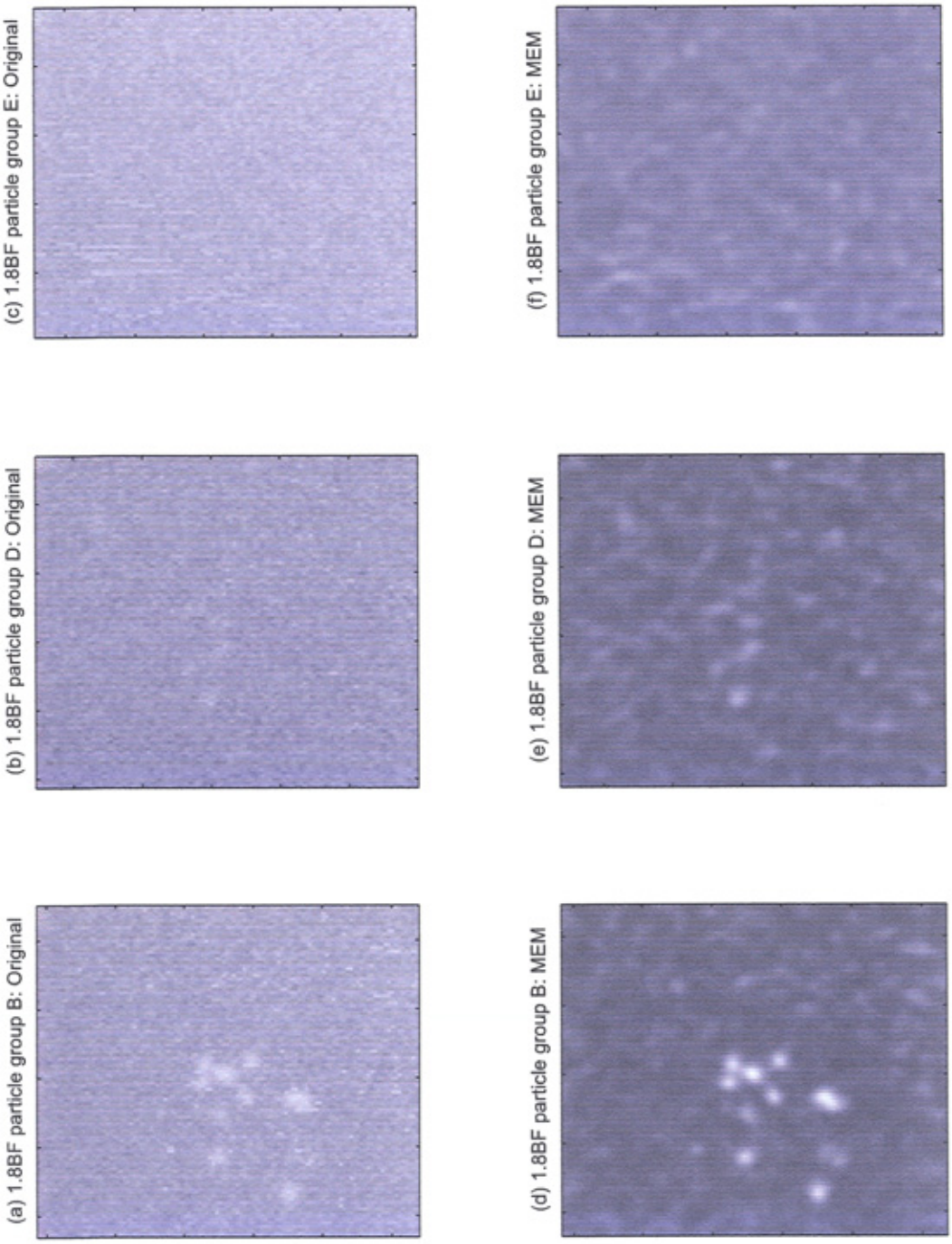

Figure 4.20: 1.8BF. Close-ups of particle groups B, D and E are shown in (a), (b) and (c) respectively. MEM restorations are shown beneath in (d), (e) and (f); improvements in visibility and resolution are apparent with the first two groups. Group E is not detectable in either the original or MEM processed images.

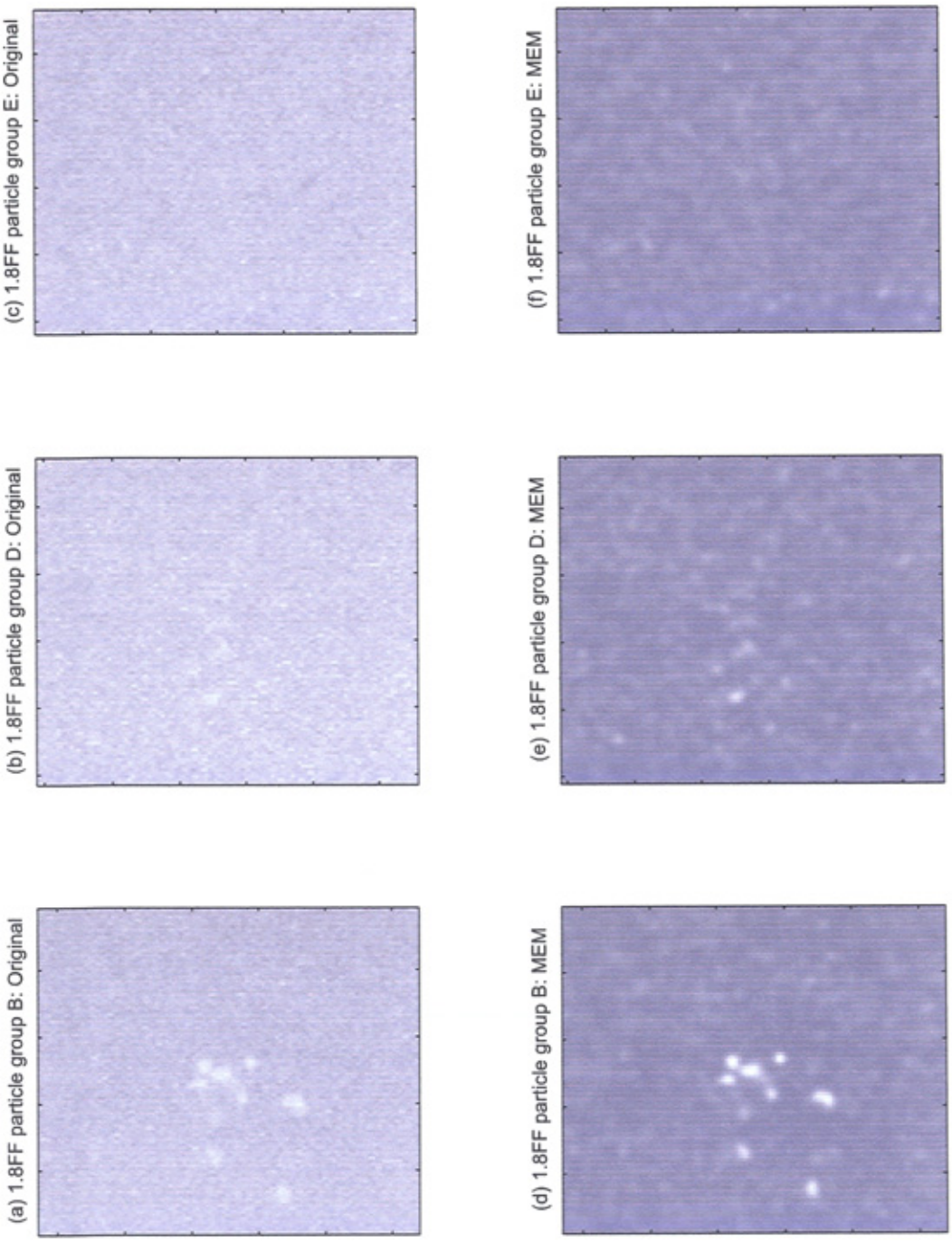

Figure 4.21: 1.8FF. Close-ups of particle groups B, D and E are shown in (a), (b) and (c) respectively. MEM restorations are shown beneath in (d), (e) and (f); improvements in visibility and resolution are apparent with the first two groups. Group E is not detectable in either the original or MEM processed images.

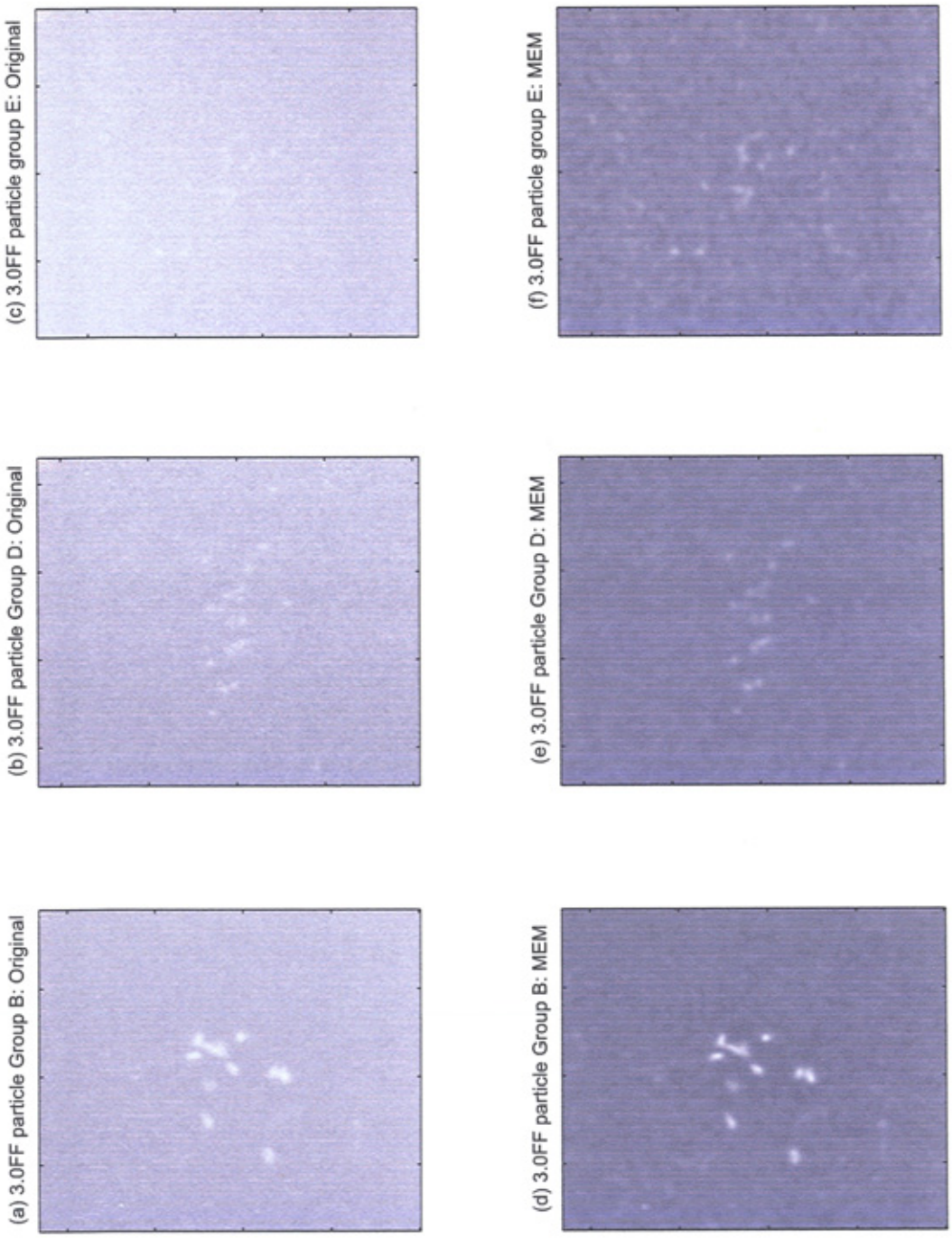

Figure 4.22: 3.0FF. Close-ups of particle groups B, D and E are shown in (a), (b) and (c) respectively. MEM restorations are shown beneath in (d), (e) and (f); improvements in visibility and resolution are apparent in all three cases, with fine details of individual microcalcifications being discernable after MEM restoration (d).

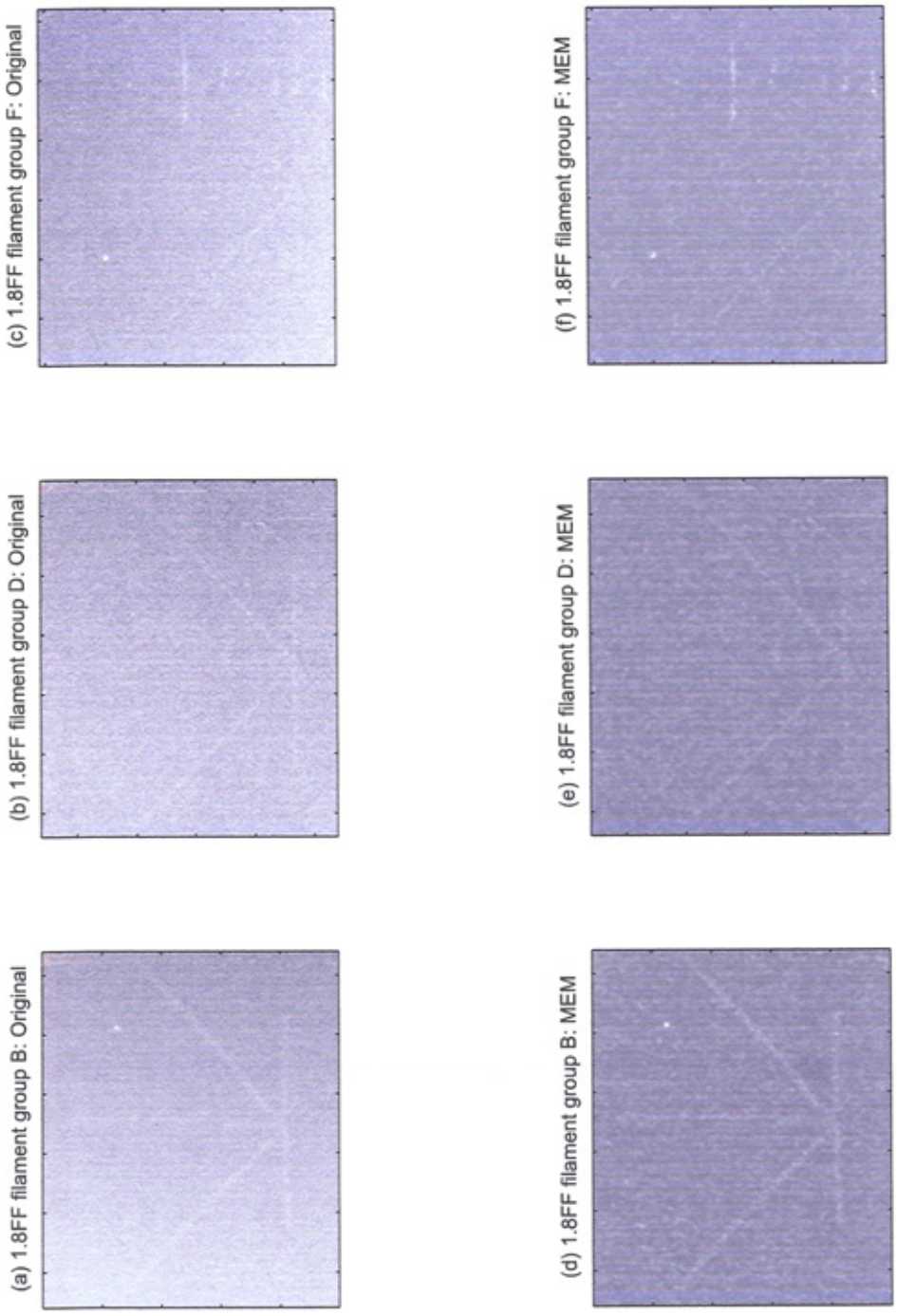

Figure 4.23: 1.8FF. Close-ups of filament groups B, D and F are shown in (a), (b) and (c) respectively. MEM restorations are shown beneath in (d), (e) and (f); improvements in visibility and resolution are apparent in all three cases.

signal-to-noise ratio perceived by the observers for the various test features, including low contrast objects, was increased by the application of the MEM processing.

Mag.	Image.	TORMAM scores			
		Filaments	Particles	Disks	Total
1.8BF	Original	20.5	6	24.5	51
	MEM	38.5	8.5	30	77
1.8FF	Original	27	6	21.5	54.5
	MEM	41	10.5	29.5	81
3.0FF	Original	42	9.5	27.5	79
	MEM	57.5	11	35.5	104

Table 4.1: Image scores of the original and MEM processed images, obtained by averaging the individual scores of the two independent observers. It should be noted that the individual scores of the two observers (not shown) were also in good agreement for the scored features.

Discussion

The aim of the work was to demonstrate an improvement in spatial resolution for realistic radiological images from MEM de-blurring, with no associated penalty in terms of reduction in the signal-to-noise ratio perceived by the observer. The expected resolution improvements are shown in figures 4.20-4.23. Somewhat unexpected, however, were the improvements in image score shown in Table 4.1, as most of the features in the TORMAM phantom, i.e. the fibre groups and disks, are essentially low-contrast features whose detection would be expected to be limited by the relative noise level in the image. The improvement in scores therefore implies an improvement in signal-to-noise ratio for this phantom. For the filaments and particles, which are comparable to the PSF in extent, improvements in visibility are effected by enhanced intensity and sharpness due to focussing, and by noise reduction. For the disks, which are significantly larger than the PSF, sharpness (resulting in easier edge detection) and noise reduction are the important factors.

It may be that the improvements in features imaged against a uniform background, as in these demonstrations, are better than those which might be achieved when imaging diagnostic features against an anatomical background. A suitable experiment to examine this will be described in section 6.2.

4.2.2 Mammographic dose reduction

The reconstructions of TORMAM images in section 4.2 were obtained using typical radiographic settings and therefore contained realistic noise levels. In this section, the quality of MEM restorations are presented from an experiment in which the conventional radiation dose is successively lowered and with corresponding detriment to the SNR in the original image (for the reasons given in section 1.3.3). The purpose of the experiment is to explore the usefulness of MEM in a noisy radiological environment and to show that image deconvolution can maintain an acceptable image quality when the radiation dose is decreased.

Experimental method

The Leeds TORMAM phantom[††] was imaged on a hospital-based Siemens Nova 3000 mammography unit. For this experiment the images were obtained under the conventional set-up (1.8× magnification using a fine focal spot). The measured fine focal spot size of this mammography unit is 0.2×0.2mm. In all cases the phantom was imaged on top of a thick stack of Perspex to provide realistic scatter and attenuation as recommended in the instructions for the phantom. The radiographic factors used were 28kVp, with successively lowering doses of 50mAs, 40mAs, 32mAs, 20mAs and 10mAs. These factors were chosen to cover values which decreased image quality compared to those used in routine mammographic quality assurance tests. The upper limit of 50mAs was determined by the Automatic Exposure Control (see page 25) for this unit. The level of quantum noise in the test images therefore ranged from realistic to very noisy. The image receptor was a Philips ACR-3 computed radiography system with a cassette having dimensions 18cm×24cm and a pixel size of 101μm (Type C on page 14).

The original and processed sets of images were viewed and scored by two independent observers both experienced in the use of mammographic image quality test phantoms. The test images were graded using the 3, 2, 1, 0 scoring system recommended in the TORMAM phantom instructions and adopted in surveys of mammographic image quality in the UK Breast Screening Programme (Young and Ramsdale 2003).

The PSF for this experiment was obtained by placing a mammography multiple pinhole test tool[‡‡] in the same plane as the TORMAM test object

[††]This was a different object to the one used in the previous experiment (as is evidenced by the different layout of the particles in the groups shown in later figures.)

[‡‡]Model M300 with 50μm pinhole size. Standard Imaging Inc., Medical College of Wisconsin, USA.

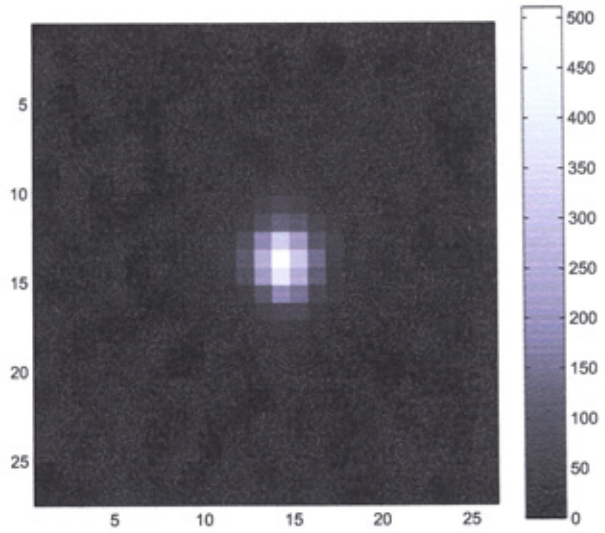

Figure 4.24: Raw image of the focal spot of the Siemens Nova 3000 mammography unit obtained at 1.8FF using a pinhole camera setup.

(unlike previous experiments, the pinhole was imaged separately from the object). The raw PSF obtained with this configuration is shown in figure 4.24.

Image processing procedure

The images obtained from the CR reader were in Siemens SPI file format which were 12-bit grayscale images holding 10-bit image data. Matlab R12 was used to read these images and perform some preprocessing so that the subsequent MEM processing could be applied. These steps were essentially the same as in the previous experiment (see page 113 for details).

MEMSYS5 was used to treat the images following the Matlab preprocessing. The default image model m was defined to be a flat image with low pixel luminance values ($m_i = 0.001$ for all i). This codifies our belief that, in the absence of data, very high photon counts were recorded (i.e. no absorption due to intervening material). Use of the program's Historic MEM switch and setting an ICF (of width 1) was done to implement the Hybrid MEM scheme of section 3.4.

The processed images contained about 3.5 million pixels. MEMSYS5 typically converged to a solution within 20 α-iterations with a processing time of 8-9 minutes, for an image of the full TORMAM test object. The smaller images presented in figures 4.25 to 4.27 were cropped from such an

Dose	Image	TORMAM scores			
		Filaments	Particles	Disks	Total
50mAs	Original	29	8	27.5	64.5
	MEM	46.5	10.5	33	90
40mAs	Original	26	7	21	54
	MEM	45.5	11	31	87.5
32mAs	Original	29	7	20.5	56.5
	MEM	43	10	30.5	83.5
20mAs	Original	16.5	7.5	19.5	43.5
	MEM	31.5	9.5	28	69
10mAs	Original	12.5	4.5	17	34
	MEM	13	6	18	37.5

Table 4.2: Image scores of the original and MEM processed images, obtained by averaging the individual scores of the two independent observers.

image after processing, rather than processed individually. The processed output files from MEMSYS5 were 8-bit PNG files. All image processing was carried out on a Pentium 4 2.4GHz machine with 512MB of RAM.

Results

Before proceeding to a systematic comparison of the original and MEM processed images and related scores, we will say a few words about our aspirations for this experiment: these were to show that the improvements to image quality could compensate (to some extent) the noise degradation introduced by using a lower x-ray dose.

The results of scoring the original and MEM processed images is shown in Table 4.2. Processing with MEM led to an improved score for all particle, filament and contrast groups at each dose setting over the original image. The degree of improvement is marginal for the images obtained at 10mAs.

Figure 4.25 shows particle group B imaged at (a) 40mAs, (c) 32mAs and (e) 20mAs. The maximum entropy reconstructions are shown alongside each image. Using pixel intensity profiles through the particles, as we did in the previous experiment (see page 116), we have calculated SNR values for each particle group image. The quoted SNR is the difference between the largest signal and the mean background within each group, divided by the standard deviation of the signal in a region close to the group. For the images shown in the figure we list SNR values for each dose. 40mAs: $SNR_{orig} = 9.2$, $SNR_{MEM} = 45.8$. 32mAs: $SNR_{orig} = 8.6$, $SNR_{MEM} = 37.0$. 20mAs: $SNR_{orig} = 6.4$, $SNR_{MEM} = 17.5$. In addition to the images shown in this

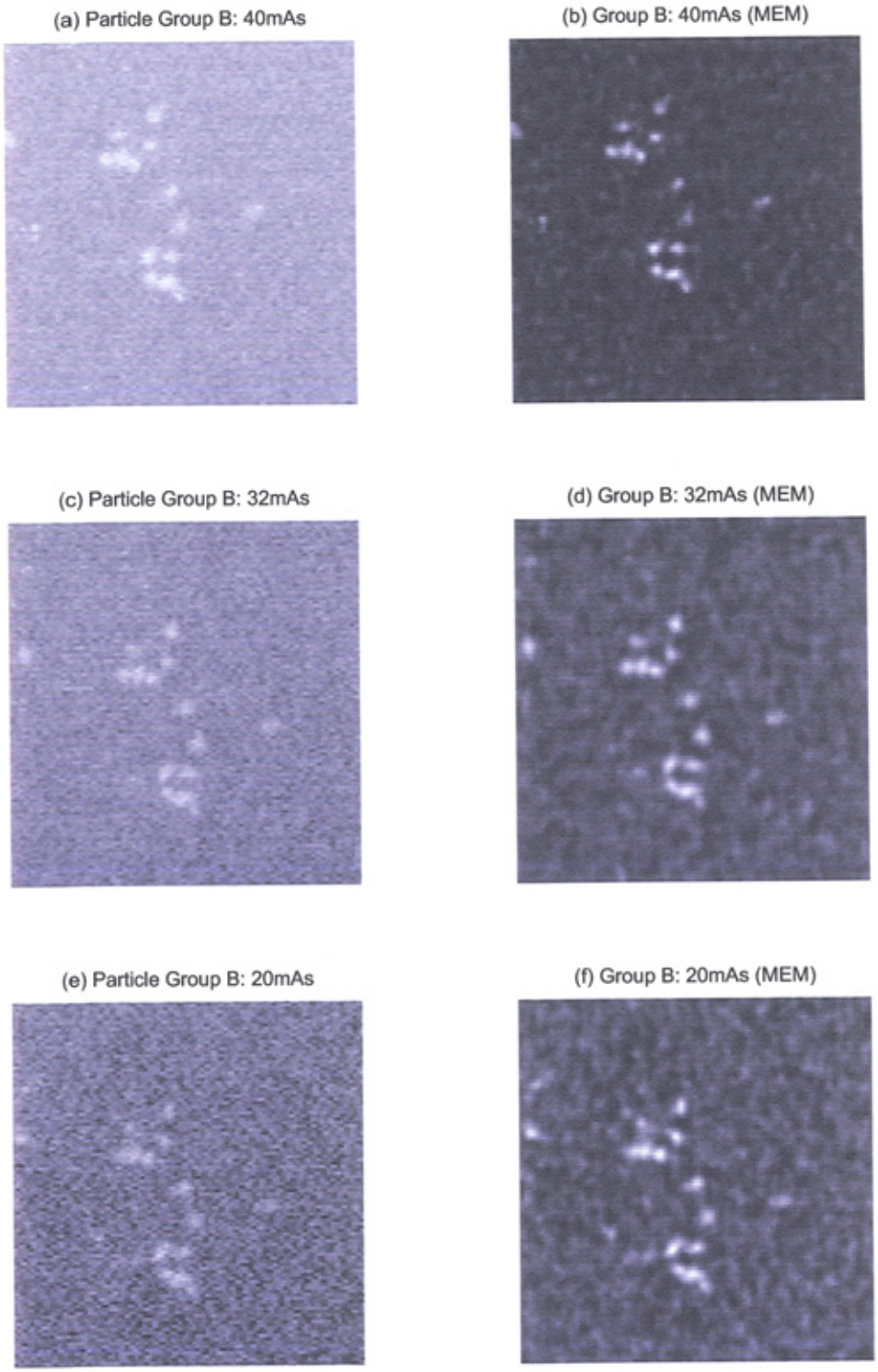

Figure 4.25: TORMAM phantom: Comparison of particle group B imaged at 40mAs, 32mAs and 20mAs with MEM restorations shown alongside each image.

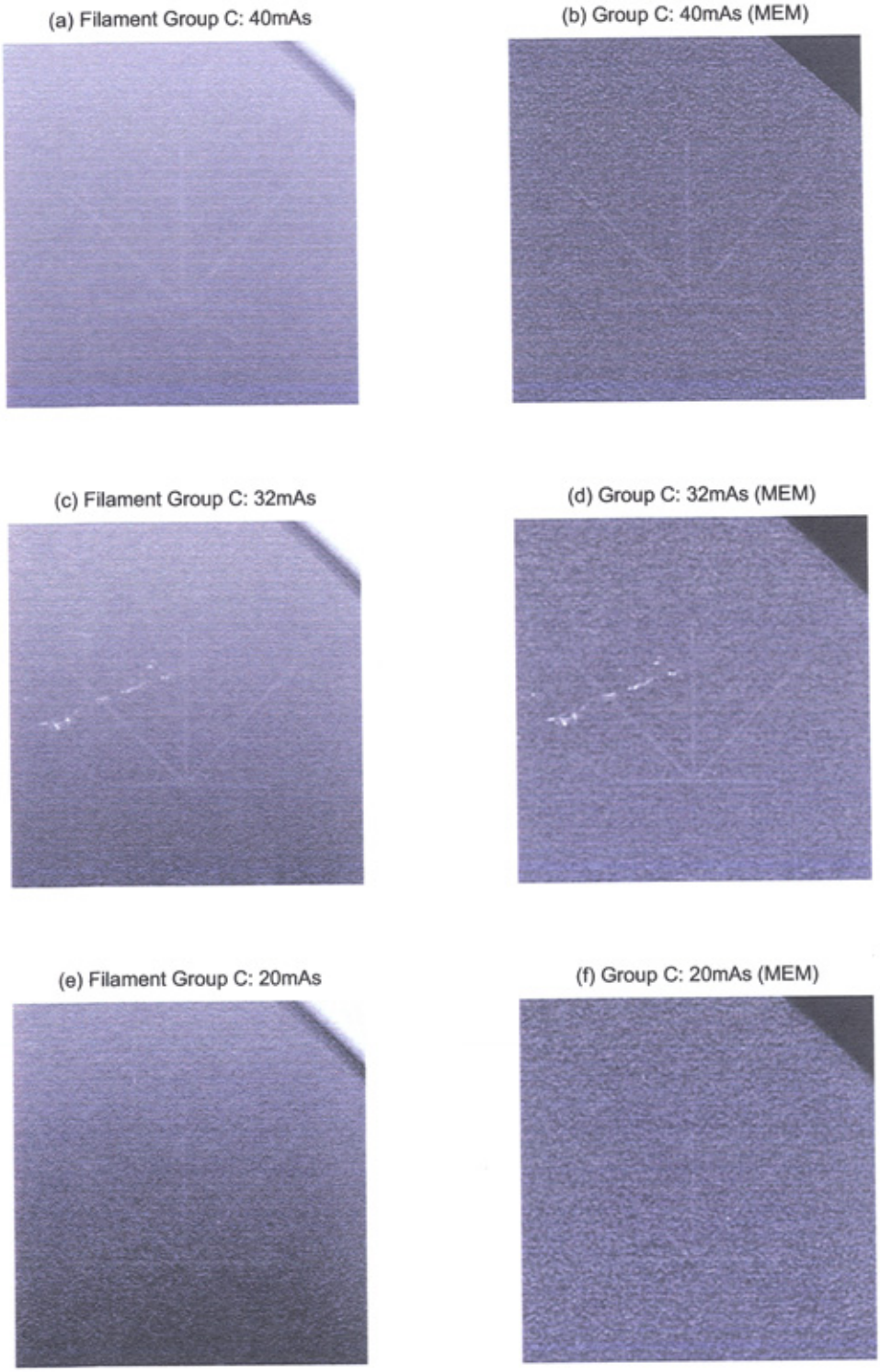

Figure 4.26: TORMAM phantom: Comparison of filament group C imaged at 40mAs, 32mAs and 20mAs with MEM restorations shown alongside each image.

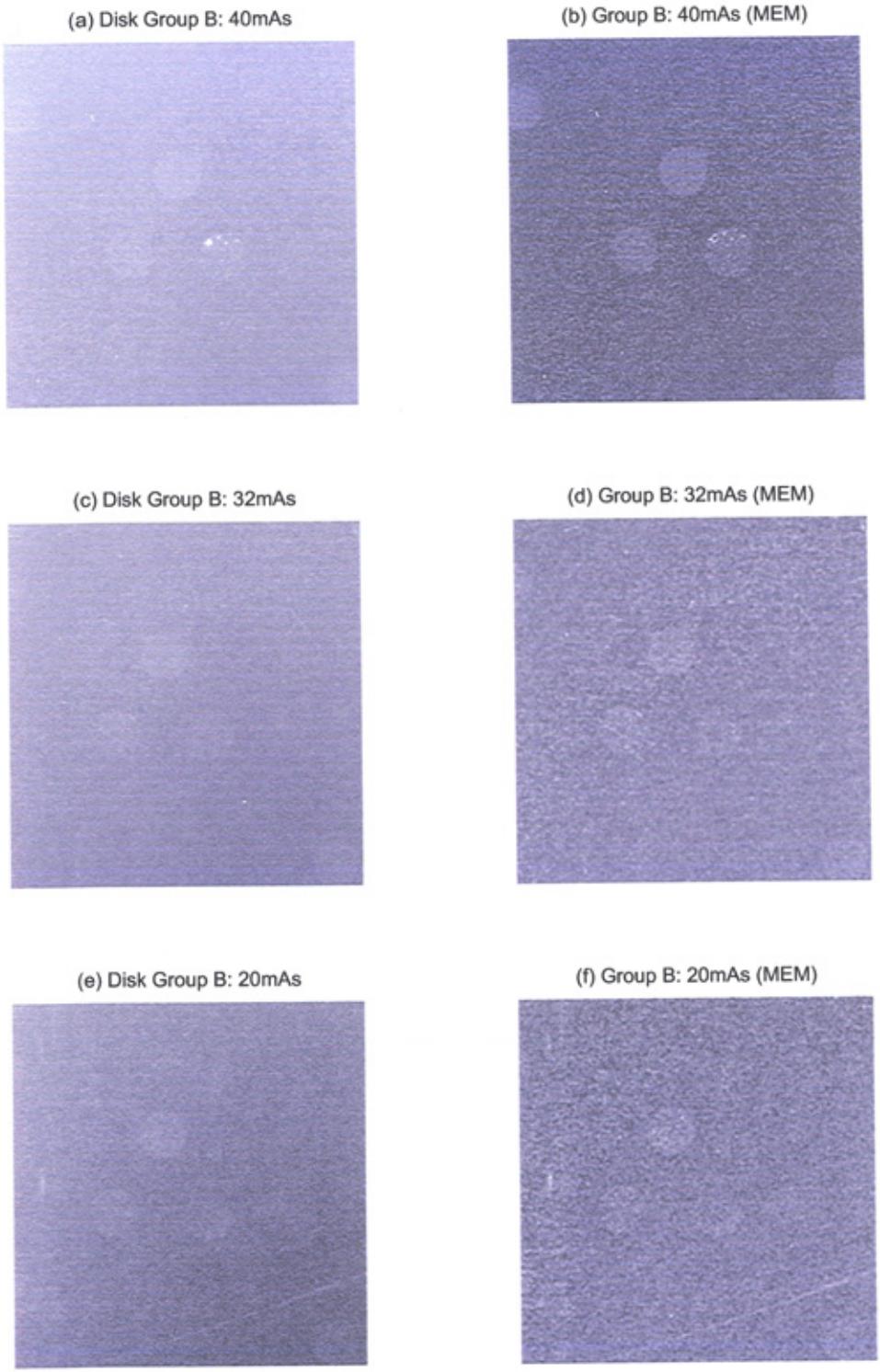

Figure 4.27: TORMAM phantom: Comparison of disk group B imaged at 40mAs, 32mAs and 20mAs with MEM restorations shown alongside each image.

figure, we also have for the lowest 10mAs dose: $SNR_{orig} = 4.1$, $SNR_{MEM} = 5.7$.

Figure 4.26 shows filament group C with the layout of original and restored images the same as in the previous figure. The visibility of the filaments remains good until noise levels prevent adequate reconstruction in the lowest dose images.

Figure 4.27 shows low contrast disk group B in the same layout as the previous figure. The visibility of these features was improved significantly in the images obtained between 32mAs to 50mAs. There is slight improvement in overall perception at 20mAs and perhaps also at 10mAs.

Discussion

The aim of this experiment was to investigate the effect of MEM on spatial resolution, feature visibility and image quality as the x-ray dose was decreased to levels much lower than those used in practice. The results indicate that a significant reduction in radiation dose is possible: the restoration process can compensate for the loss of image quality caused by increased noise levels. Inspection of Table 4.2 shows that the scores of the unprocessed 50mAs image (the dose selected by the AEC) are comparable to the scores of the 20mAs MEM image. The scores of the MEM processed 32mAs image are significantly higher. We suggest that a dose reduction from 50mAs to 40mAs or 32mAs, followed MEM processing results in better visibility of features in the TORMAM object than would be obtained from just using the higher dose setting.

Our belief is that such a performance would be of clinical interest for two reasons. Firstly, it would be beneficial if the standard x-ray dose used in mammography could be lowered provided that image processing could be shown to have no detrimental effect on image quality. Secondly, maintaining image quality at reduced doses would give a radiographer more options from which to choose radiographic settings. Examples include higher than standard magnifications (>1.8) or larger focal spots used in conjunction with shorter exposures at conventional magnification. A more thorough discussion of the ramifications of dose reduction (and the other ideas from this chapter) on system optimisation will be presented in chapter 7.

CHAPTER 5

MEM applied to linear tomography

The process of eliciting depth information using linear tomography was introduced in chapter 1. The combined movements of the focal spot and image receptor (see figure 1.9 on page 27) allow a thin 'slice' of a 3-D object to be imaged.

Images obtained from linear tomography can be thought of as having two components; the first is the focal plane, which has sharp details and the features of interest in that plane. The second component is blurring caused by the apparent motion on the image receptor of features above and below the focal plane. The second component of these images reduces the contrast of the features in the focal plane. Mechanical methods to reduce the out-of-plane blurring were discussed in section 1.5.2.

In this chapter we will summarise related research and other methods of blur-reduction in tomography images. A simple, simulated linear tomography model will be explored and a modification of MEM suitable for the problem will be presented. Finally, the adapted method will be applied to images of test objects and a clinical anthropomorphic skull phantom.

5.1 Related methods and research

Image subtraction

The first attempts to tackle the problem of blurring in linear tomography images were variations on the theme of image differences. Edholm and Quiding (1969; 1970) described photographic methods, where the original film

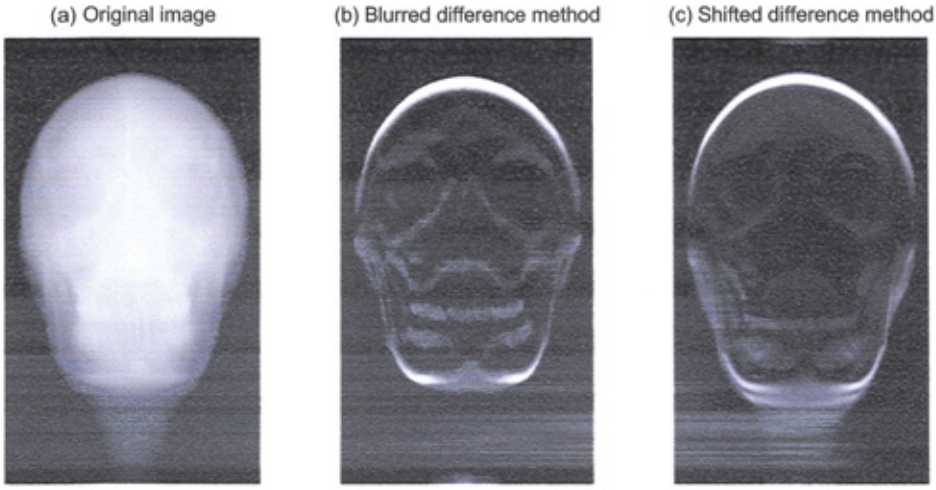

Figure 5.1: Examples of the image subtraction techniques proposed by Edholm and Quiding and Chakraborty et al. (a) is the original image of a skull phantom. (b) is the difference image formed by blurring a negative of (a) in the same direction as the linear movement by a distance of 25 pixels. (c) is the difference image formed by displacing an identical copy of (a) by 5 pixels in the direction of the linear movement.

image is superposed on a blurred negative copy. The authors also described a system with a video camera, intensifier screen and television display to perform variations of image subtraction electronically and almost in real-time. These photographic and electronic techniques are essentially a form of manual high-pass filtering. An equivalent process applied to digital images was described by Chakraborty et al. (1984) and applied to digitised tomosynthesis images. Examples of these image subtraction methods applied to a tomographic image of a skull phantom are shown in figure 5.1.

A disadvantage of image subtraction methods is that they reduce or remove useful low frequency information lying within the plane of interest. Direct high pass frequency filtering will be examined more closely in section 5.1 with particular emphasis given to consideration of the direction and effects of linear blurring.

Restoration of tomosythesised images

Tomosynthesis is a technique for obtaining tomographic information through an arbitrary cross-section of an object (see page 28). Tomosythesised images suffer the same out-of-plane blurring as conventional tomographic images and

related research to reduce this blurring is now described.

A method to reduce tomosynthesis blur was developed by Ruttimann et al. (1984) and applied to images obtained with circular motion of the focal spot and image receptor. The authors of that paper used a model which approximated the 3-D object as a series of n discrete planes which are represented by the tomosythesised images. The blurring in a particular image is assumed to be due to contributions from structures within the other $n-1$ planes. This formulation leads to a system of coupled linear equations. Weighted estimates of the 'true' planes were obtained and convolved with blurring functions derived from a consideration of the geometry of the imaging system. The in-plane blurring was calculated iteratively and subtracted from the original images to increase the visibility of features in each focal plane.

Using a similar image model Ghosh Roy et al. (1985) attempted to remove blur, due to circular motion, from a particular tomosynthesised plane. The blur was modeled as being caused by contributions from several adjacent planes in the object. Solutions to the resulting system of linear equations were sought to obtain reconstructed planes without the blur. Matrix inversion tomosynthesis (MITS) was an extension of this work, developed by Dobbins et al. (1987). MITS is a technique which attempts to exactly solve the coupled system of linear equations by direct inversion of a matrix of blurring functions. Although MITS is computationally fast and fairly successful and removing blur the method is susceptible to noise at low frequencies.

Wavelet based approaches to restoration

Wavelet based removal of blur and/or noise from digital tomography (and tomosynthesis) images is described by several investigators. The approach of Badea et al. (1998) exploits a specific noise pattern within tomosynthesised images and uses the spatial locality of the wavelet transform to discriminate between features inside and outside the fulcrum plane.

Wavelet analysis has also been applied to denoising images obtained with limited angle tomography (Sahiner and Yagle 1993) and fan-beam tomography (Bonnet et al. 2002). Wavelet-based Bayesian estimators for removal of Poisson noise from emission and transmission tomography images is discussed by Huang et al. (2003).

Improved frequency filtering

The partial success of the image subtraction approaches of section 5.1 can be attributed to its equivalency to high-pass frequency filtering. It is possible

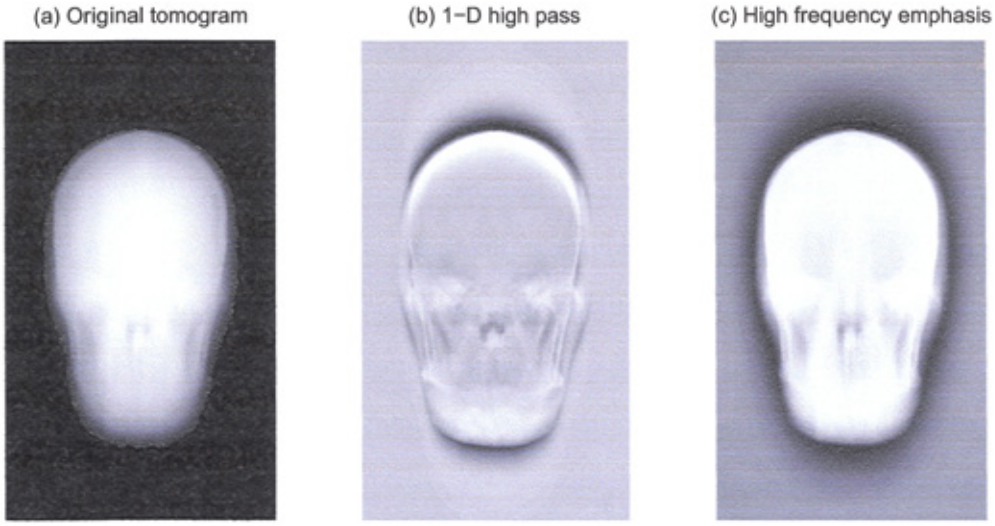

Figure 5.2: (a) Tomogram of a skull phantom. (b) Application of a 1D high-pass filter. (c) Result of applying a high frequency emphasis filter to the original image (parameters used: a=0.5, b=2.0)

to design frequency filters specifically for the case of images blurred by linear motion. Since the blurring occurs in only one direction, a 1-dimensional high-pass filter can be constructed to suppress those frequencies associated with directional blurring. A tomogram of a skull phantom with the associated Fourier transform was shown in figure 2.1 (on page 34). It was seen that the low frequency content was greatly increased in the direction associated with the linear blur.

Figure 5.2b shows the result of applying a 1-D filter which aims to ensure that frequencies associated with real features and low frequency content in a direction perpendicular to the linear blurring are retained.

High frequency emphasis filtering

The fine structure and details of an image d are located at high frequencies in the Fourier transform of the image D. A useful method of enhancing radiographic images is the high-frequency emphasis (HFE) filter (Gonzalez and Woods 2001).

The first step is to construct a high pass filter H_{uv}. A cut-off frequency is chosen and the high pass filter is formed by retaining only those Fourier components higher than the cut-off frequency and setting those below it to zero. The HFE filter $\hat{H}_{uv}$ is created by modifying the high pass filter:

$$\hat{H}_{uv} = a + bH_{uv} \tag{5.1}$$

where a, b are constants and $a \geq 0$ and $b > a$. The filtered image is formed by multiplying the Fourier transform of the HFE filter (equation 5.1) with the Fourier transform of the original image:

$$\hat{D}_{uv} = \hat{H}_{uv} D_{uv} \tag{5.2}$$

This method can be better than a straightforward high pass filter because lower frequencies (larger structures) are also retained. HFE filtering can be used to bring out the edges and sharp transitions which typify features close to the focal plane in tomographic images. The Matlab code for our HFE filter is listed in Appendix D.2.

The clear disadvantage of this method is that it does not distinguish between noise and signal; both will be enhanced. The method is more suited to emphasising edges than the detection of small microcalcifications. Also any large scale, low frequency features in the focal plane of the tomogram will reduced in contrast.

Expectation Maximization methods

The Expectation Maximization (EM) algorithm (Dempster et al. 1977) is a method used to approximate a PDF. Since an image can be thought of as a PDF (see page 68) the EM algorithm can be applied to some image restoration problems. EM is typically used in situations where we have incomplete sampling of the true object distribution: this is the case in tomography. EM is used to compute maximum likelihood estimates of the missing samples. In the astronomical community EM is essentially the same technique as the method known as Richardson-Lucy deconvolution (see section 2.3.6).

EM has been widely applied to finding solutions which maximise a likelihood function for quantum noise limited tomographic data. One of the first applications was to emission tomography by Shepp and Vardi (1982). Following that work a number of researchers applied the method to medical imaging problems. However, widespread acceptance of EM suffered until image processing times, which were at first too great to be of clinical use, and the problematic presence of noise artefacts in restored images were addressed (Miller and Roysam 1991). Development and application of EM methods is an area of ongoing research. For examples of notable applications in other areas of tomographic imaging see Hudson and Larkin (1994), De Pierro (1995) and Kinahan et al. (1997).

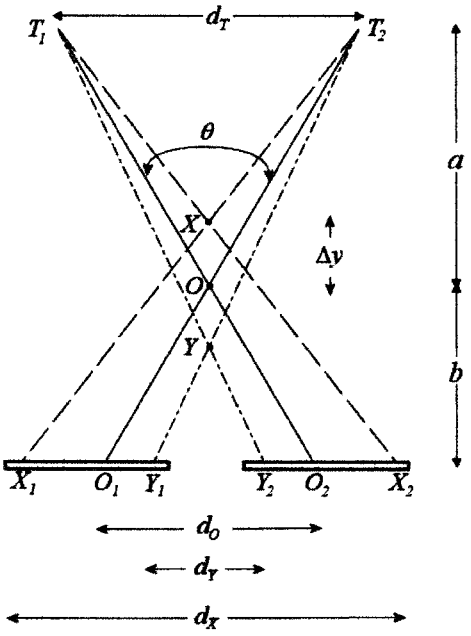

Figure 5.3: Layout of a linear tomography imaging system showing the distances necessary to calculate blurring on the image receptor.

5.2 Modeling the linear tomography system

5.2.1 Calculation of tomographic blurring

Figure 5.3 shows the geometry of a linear tomography system. The x-ray tube moves from position T_1 through a distance d_T to position T_2. The movement of the tube would cause the projected shadows of points X, O and Y in the object to become smeared on the image receptor by different degrees (d_X, d_O and d_Y respectively) if the receptor was fixed. However, the image receptor is designed moved in an opposing direction to the tube in order to keep projected points in same plane as O (the focal plane) from becoming blurred on the receptor. Therefore, if the image receptor is moved through a distance d_O then the shadow of point O will not become blurred. The distance that the image receptor needs to move to achieve this is given by:

$$d_O = d_T \times \frac{b}{a} \tag{5.3}$$

Suppose the point X lies some distance Δy above O. The corresponding projected shadow traverses a distance on a stationary image receptor given by:

$$d_X = d_T \times \frac{b + \Delta y}{a - \Delta y} \tag{5.4}$$

Now consider the situation in which the image receptor is moved to make the plane containing O the focal plane. The movement of X's shadow relative to the image receptor will describe the amount of tomographic blur B_X for the plane containing X. This is given by:

$$\begin{aligned} B_X &= d_X - d_O \qquad (5.5) \\ &= d_T \left(\frac{b + \Delta y}{a - \Delta y} - \frac{b}{a} \right) \\ &= d_T \left[\frac{a(b + \Delta y) - b(a - \Delta y)}{a(a - \Delta y)} \right] \\ B_X &= \frac{d_T \Delta y}{a} \left(\frac{a + b}{a - \Delta y} \right) \qquad (5.6) \end{aligned}$$

For our experiments in this work it is desirable to express the out-of-focal plane blur in terms of the tomographic swing angle θ rather the distance d_T. It is simple to show that $d_T = 2a \tan(\theta/2)$, so we can rewrite equation 5.6 as:

$$B_X = 2\Delta y \tan\left(\frac{\theta}{2}\right) \left(\frac{a + b}{a - \Delta y} \right) \tag{5.7}$$

5.2.2 Adapting MEM to linear tomography

Simulated linear tomography images

To illustrate the problem and to formulate a solution I have simulated a simple test object comprising the numbers '3', '6' and '9' situated at different increasing depths. Commercial graphics editing software* was used to create an 8-bit grayscale basis image, and to 'motion blur' some of the components by appropriate amounts in a horizontal direction. Blurring functions of 15 and 30 pixels in length were used. These images were further degraded with a small amount of noise using Matlab: Gaussian noise with zero mean

*Paint Shop Pro 8: Jasc Software

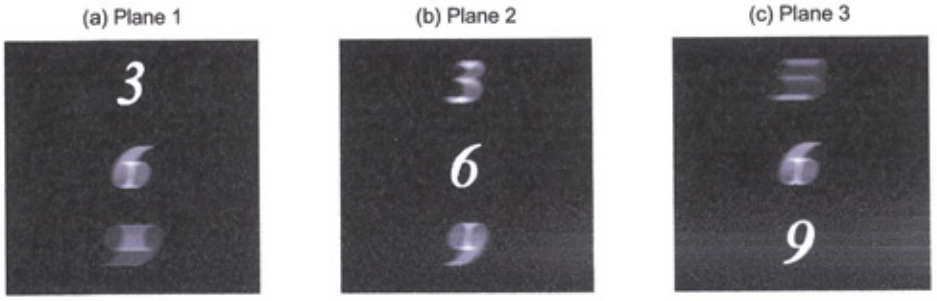

Figure 5.4: Images from a simulated linear tomography system. Three objects are imaged three times, with the focal plane passing through a different object each time. The planes outside the focal plane are blurred by either 15 or 30 pixels, depending on their supposed distance from the focal plane

and variance $\sigma^2 = 0.0001$ was added to each. The three simulated linear tomography images shown in figure 5.4.

We denote the corresponding true images by x_1, x_2 and x_3 respectively. The images d_1, d_2 and d_3 recorded by a linear tomography machine are combinations of true images degraded by PSFs, dependent on the geometry of the machine and distances between the object, focal spot and image receptor. Figure 5.4 represents three scans of a test object. In the first row the focal plane is close to '3', '6' in the second and '9' in the third. In the first case the recorded image is:

$$d_1 = A_{11} * x_1 + A_{12} * x_2 + A_{13} * x_3 \tag{5.8}$$

where A_{11}, A_{12} and A_{13} are linear point-spread functions corresponding to the respective depths. Similarly we have scans 2 and 3 with respectively deeper focal planes and different point-spread functions, and equation (5.8) generalises to:

$$d_i = \sum_{j=1}^{3} A_{ij} * x_j \quad i = 1 \; to \; 3 \tag{5.9}$$

which is the forward map in this case. Note that the sum here is over different images, not elements of one image. The three recorded images are shown in figure 5.4 with a pictorial representation of equation (4.10) in each case.

We use the Hybrid MEM scheme (section 3.4) as the basis for our tomographic restoration algorithm. The entropy of the true image x_1 is:

$$S_1 = \sum [x_1 - m_1 - x_1 \ln(x_1/m_1)] \tag{5.10}$$

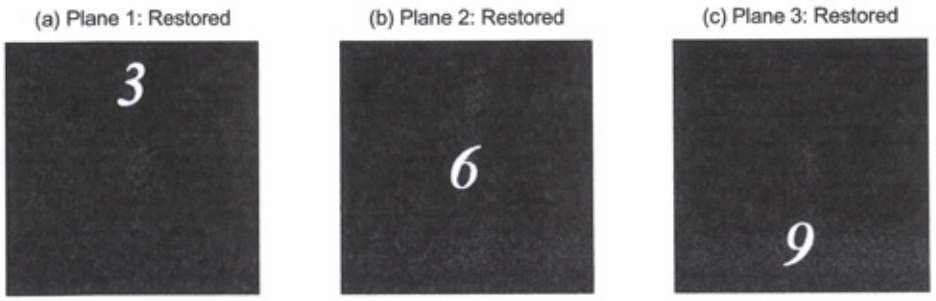

Figure 5.5: Restoration of the simulated linear tomography images shown in figure 5.4. The blurred images caused by objects outside the focal plane have almost been eliminated.

The entropy of the three images together is:

$$S = \sum_{j=1}^{3} S_j \tag{5.11}$$

The problem is to reconstruct the true images x_1, x_2 and x_3 from the recorded images. We assumed that the true images are confined to the chosen focal planes. The χ^2 statistic describing the goodness-of-fit between the degraded images and the postulated true images is given by:

$$\chi^2 = \frac{\sum_{i=1}^{3} \left(\mathbf{d}_i - \sum_{j=1}^{3} A_{ij} * \mathbf{x}_j \right)^2}{\sigma^2} \tag{5.12}$$

where σ^2 is the noise contained in each pixel, in this case assumed to be constant across the three images. Having defined the entropy and the χ^2 statistic for this situation, the procedure is as outlined in section 4.1: the problem to maximise the entropy with respect to the elements of the three vectors u_i subject to $\chi^2 = N$ as a constraint, where N is now the total number of pixels in the recorded images.

Results of the simulated tomogram restoration

Figure 5.5 shows the results of applying the modified MEM algorithm to the simulated tomograms. The out-of-focal plane blurring has been greatly suppressed in each case.

The purpose of this simple exercise was to develop and test the modified MEM algorithm under ideal circumstances; pre-determined blurring functions, little or no noise and no background or bad pixels.

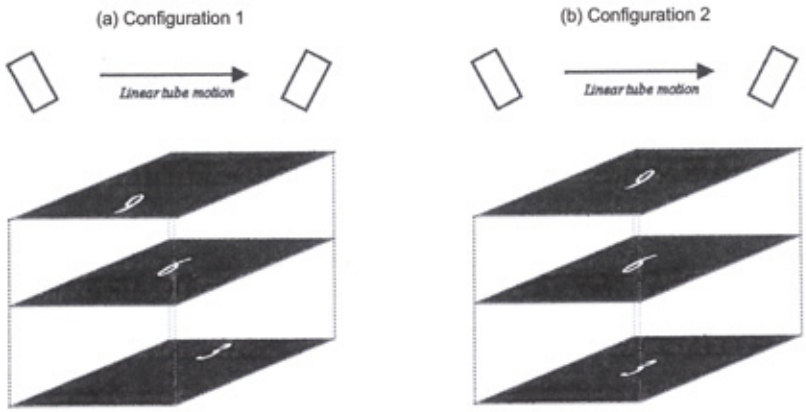

Figure 5.6: Illustration of the two configurations for imaging the lead number test object. In (a) the lead numbers are adjacent on separate planes whereas in (b) they are directly above one another giving rise to overlapping blur on the image receptor

5.3 Experiments with a 3-plane test object

A simple test object comprising three sheets of perspex was constructed. Between each layer of perspex was placed an object: a lead number of negligible thickness.

5.3.1 Method

The test object was placed on top of a table and imaged by Philips Bucky Diagnost system with linear tomography option. The focal spot to image receptor distance was 100cm and the table-top was 6.5cm above the image receptor. Cut heights through the object (measured from the table-top) of 4.2cm, 8.4cm and 12.7cm were obtained. The radiographic factors used were 70kVp, 32mA with an exposure time of 1.2s. These factors were chosen to give realistic levels of quantum noise in the resulting images. The image receptor was a Philips ACR-3 computed radiography system. The two types of film cassettes used in these experiments have resolutions of $R = 50.3$ pixels/cm and $R = 65.7$ pixels/cm respectively.

The test object was imaged with the lead numbers in two configurations. These configurations are shown in figure 5.6. In the first configuration the numbers were adjacent each other so that no overlap would occur on the image receptor. The receptor resolution in this configuration was $R = 50.3$

pixels/cm. The resulting tomographic images from this set-up are shown in figure 5.8. In the second configuration, all three objects were above one another, but separated by layers of perspex. The receptor resolution in this configuration was $R = 65.7$ pixels/cm. This set-up leads to overlapping blurring in the captured image, as can be seen in figure 5.10.

5.3.2 Calculation of blurring functions

The implicit assumption of these calculations is that geometric unsharpness caused by the focal spot size is negligible. The quantity given by equation 5.7 is a distance on the image receptor. On a digital image receptor (or after the image has been digitised by a CR reader) the length of the blurring, measured in pixels, depends on the resolution R of the image receptor.

$$B_{\text{pixels}} = 2\Delta y \tan\left(\frac{\theta}{2}\right)\left(\frac{a+b}{a-\Delta y}\right) \times R \tag{5.13}$$

As an example we calculate the length of one blurring function. Consider the tomograms through Plane 1, shown in figures 5.8a and 5.10a. What is the length (in pixels) of the blurring function which has smeared out the lead number from Plane 2?

Using the notation of section 5.2.1 we have:

$$\begin{aligned} b &= 6.5 + 4.2 = 10.7\text{cm} \\ a &= 100 - 10.7 = 89.3\text{cm} \\ \Delta y &= 8.4 - 4.2 = 4.2\text{cm} \\ \theta &= 40^\circ \\ R &= 50.3\text{pixels/cm} \end{aligned}$$

Substituting these values into equation 5.13 gives:

$$\begin{aligned} B_{\text{plane2}} &= 2 \times 4.2 \times \tan 20^\circ \times \left(\frac{89.3 + 10.7}{89.3 - 4.2}\right) \times 50.3 \\ B_{\text{plane2}} &= 180.7106 \end{aligned}$$

As we are dealing with discrete pixels the length must be an integer value, we round the answer obtained above to get:

$$B_{\text{plane2}} = 181\text{pixels} \tag{5.14}$$

The PSF which describes this blurring function is therefore a vector† with 181 elements whose values are of the form:

$$\mathrm{PSF}_{\mathrm{plane2}} = \frac{1}{181}[1, 1 \ldots, 1] \tag{5.15}$$

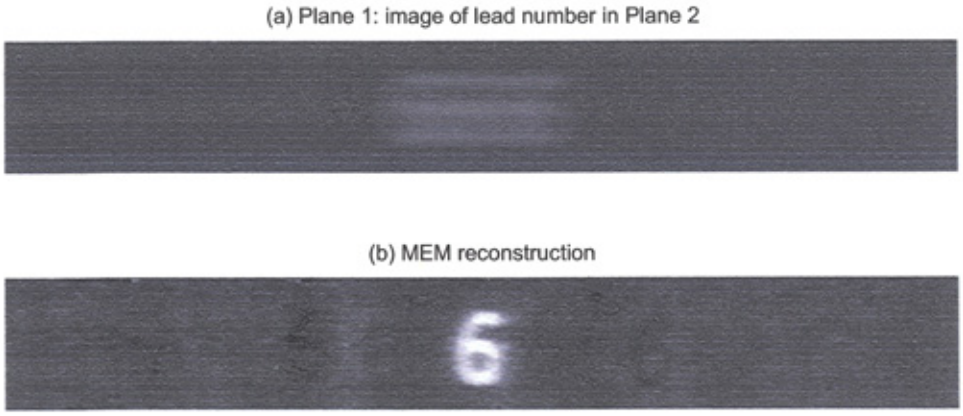

Figure 5.7: (a) shows a region of the image focussing on Plane 1 with only the image of the object in Plane 2 visible. The blurring function for this object is calculated as an example in the text. (b) shows the result of straightforward deconvolution with MEM using this blurring function. The object in Plane 2 has been recovered.

We can verify that this is correct by deconvolving just the image of the lead number in Plane 2 with the calculated PSF. The result of this test example is shown in figure 5.7. The relative lengths of the PSFs for each plane are given in Table 5.1. The actual lengths on the receptor, measured in pixels, is obviously dependent on the resolution of a particular receptor; they are calculated by multiplying the relative values by the R and rounding the result. The object in the focal plane is unblurred - this is represented as a convolution with a unit impulse (a PSF of unit length) and is independent of resolution.

5.3.3 Image processing procedure

The test-object images were pre-processed using Matlab. The image data were 'cleaned' to remove any spurious bright pixels (see page 98) and lin-

†The images considered in this chapter were subject to blurring in a vertical or horizontal direction (depending on the orientation of the image receptor) and so blurring functions are either column vectors or row vectors

	Plane 1	Plane 2	Plane 3
Plane 1	1	3.6R	7.7R
Plane 2	3.4R	1	3.8R
Plane 3	6.9R	3.7R	1

Table 5.1: The relative lengths of the blurring functions present in each plane of the lead-number test object. R is receptor resolution.

earised via equation (4.4) as required by our MEM algorithm. It is worth noting another convention in dealing with this particular restoration problem. In forming 'trial' restorations the estimates of the true object distribution are convolved with various PSFs. In this experiment the convolution operation requires significant padding of the image. Fictitious pixel values beyond the image border were taken to be equal to pixel values along the image border. The Matlab code for this 3-plane reconstruction is given in Appendix D.6. The final reconstructions were 'nonlinearised' to match the form in which the original images were displayed.

5.3.4 Discussion of results

The three reconstructed planes of configuration 1 (the adjacent lead numbers) is shown in figure 5.9. Results obtained from Configuration 2 (overlapping lead numbers) are shown in figure 5.11.

The restoration of the lead numbers in a non-overlapping configuration produces the best results. This is to be expected but is not the most realistic situation. The restoration of lead numbers whose images overlap on the image receptor also does reasonably well but artefacts caused by 'lost' information are obvious.

The experiments conducted using this phantom were useful for testing and optimizing the modified MEM code in Matlab. However, the object is not particularly realistic: the human body obviously is not comprised of discrete slices. In chapter 6 we describe an experiment using tomographic images of a more realistic 3-D object.

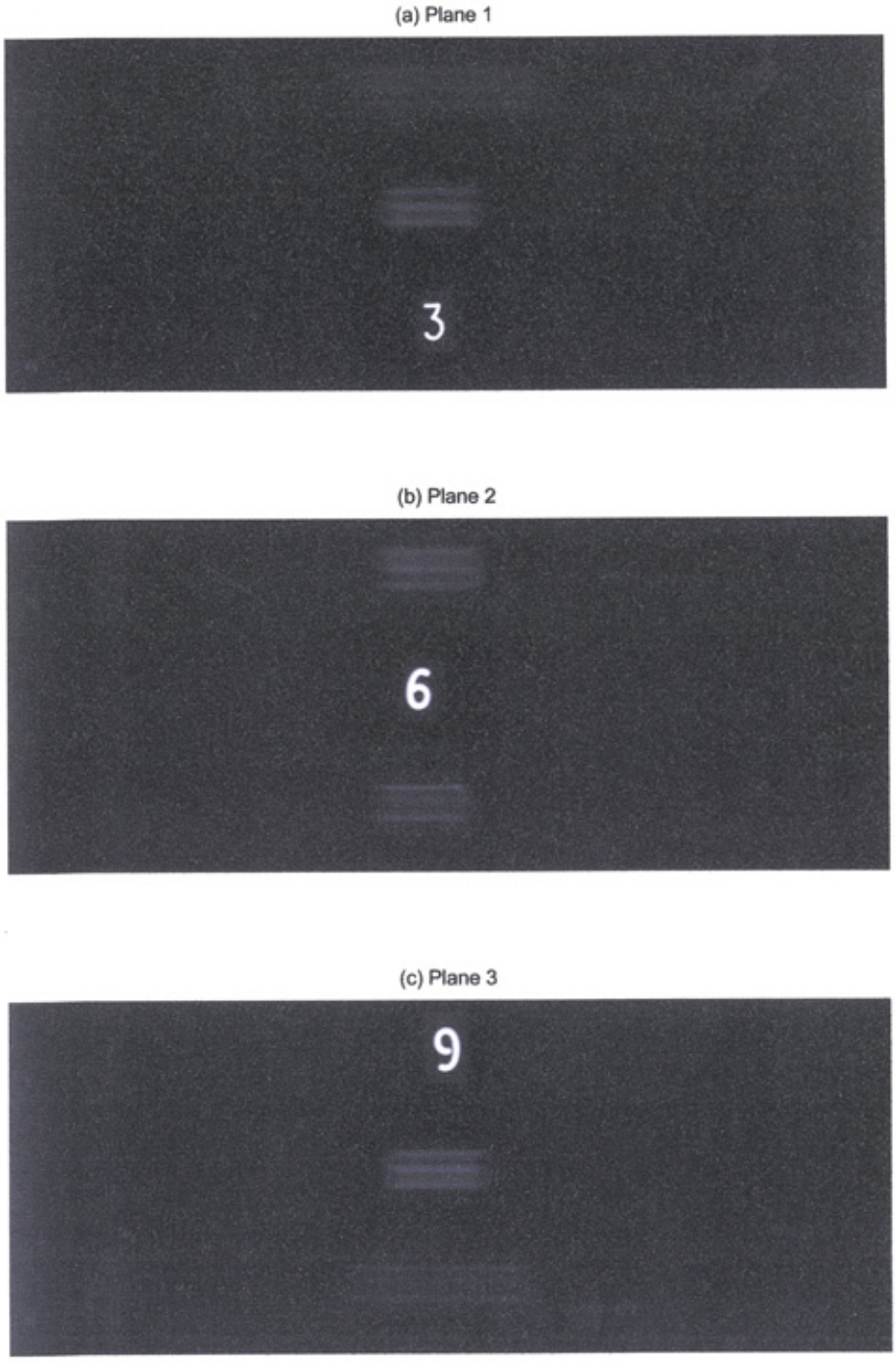

Figure 5.8: Original tomograms of the lead number test object using configuration 1. (a) Cut height 4.2cm (b) Cut height 8.4cm (c) Cut height 12.7cm

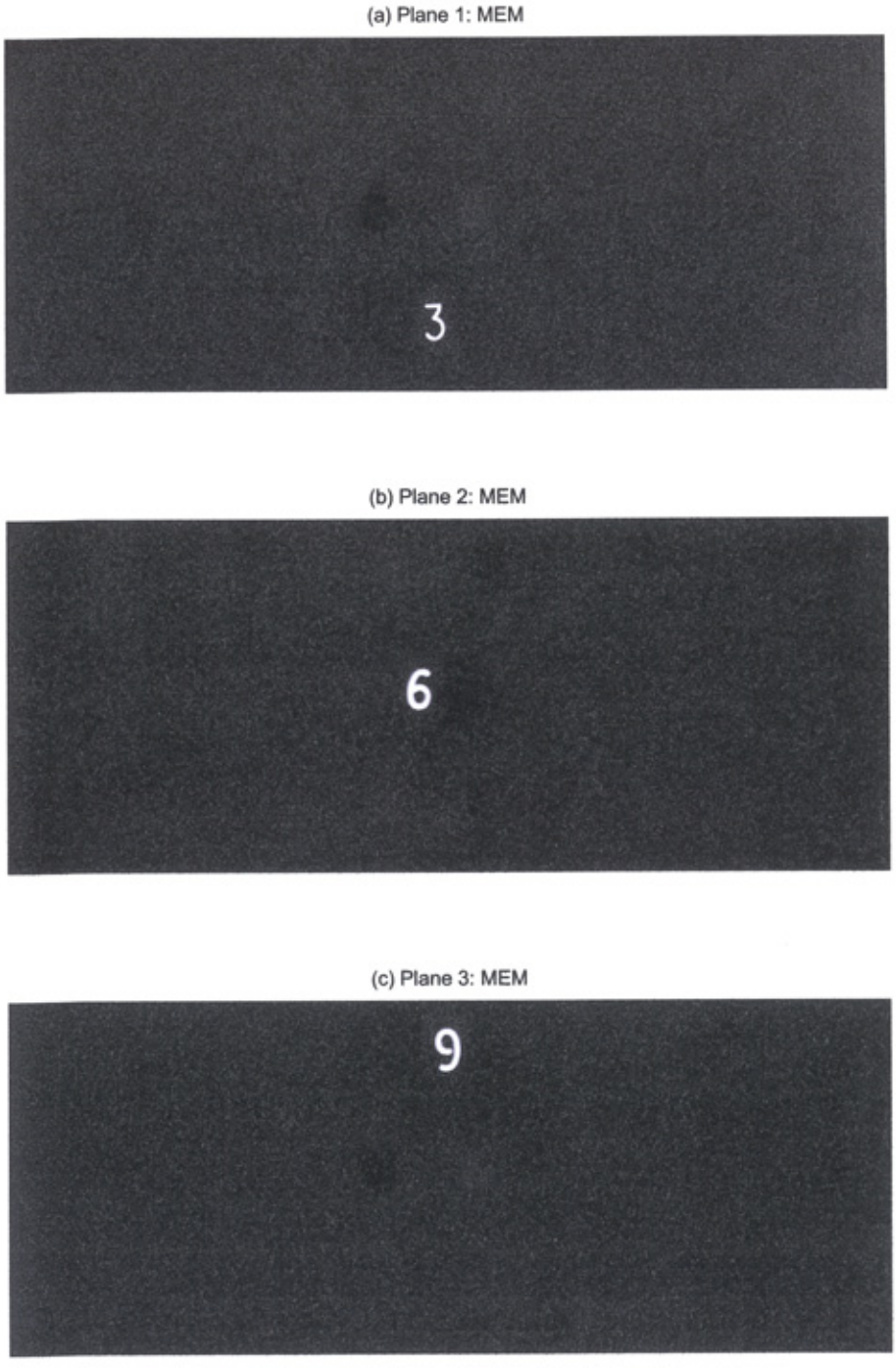

Figure 5.9: MEM reconstructions of the lead numbers arranged in configuration 1. (a) Cut height 4.2cm (b) Cut height 8.4cm (c) Cut height 12.7cm

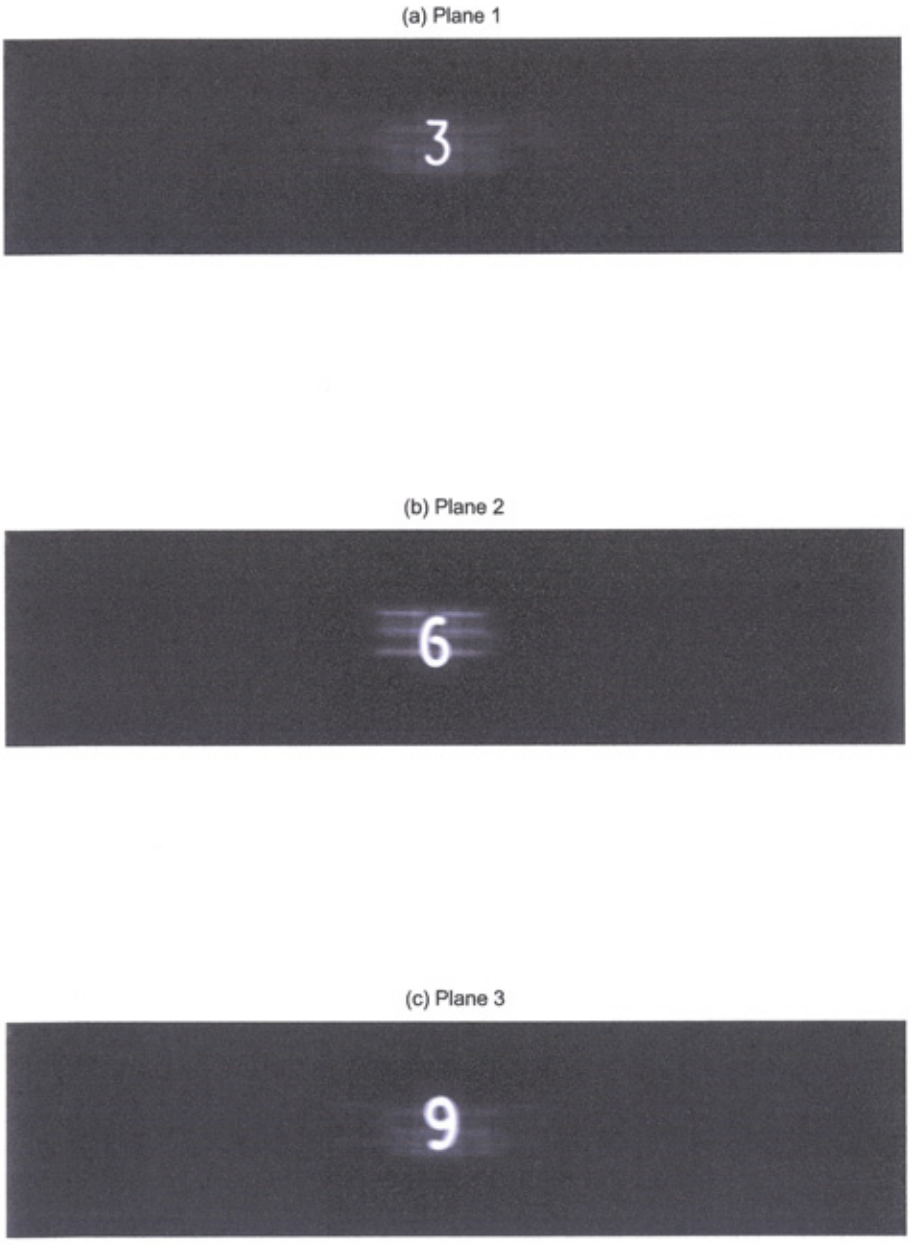

Figure 5.10: Original tomograms of the lead number test object using configuration 2. (a) Cut height 4.2cm (b) Cut height 8.4cm (c) Cut height 12.7cm

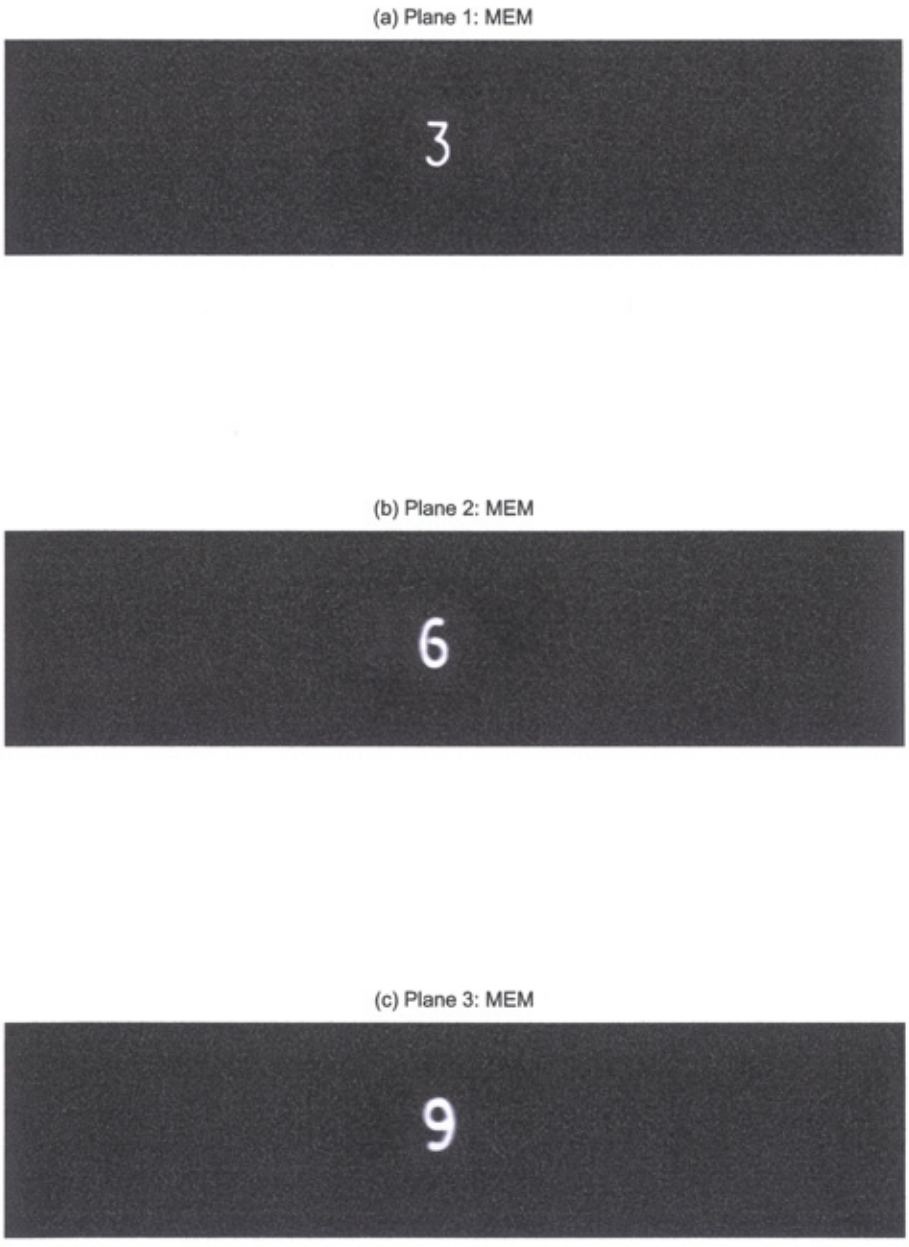

Figure 5.11: MEM reconstructions of the lead numbers arranged in configuration 2. (a) Cut height 4.2cm (b) Cut height 8.4cm (c) Cut height 12.7cm

CHAPTER 6

Application to realistic images

The experiments of the previous chapters were carried on a selection of phantoms designed to evaluate and in some cases, to quantify the success of maximum entropy deconvolution. The phantoms were unrealistic in other respects because the internal structures of the body are much more complicated. For example features of interest to a radiologist may be embedded in complicated backgrounds or hidden by other structures. In this chapter we apply our deconvolution strategies to more realistic objects and clinical data.

6.1 Magnified bone image

The images and other data used in this experiment were generously provided by Prof. Geoff Dougherty from research into the effectiveness of a commercial Wiener filter applied to radiological images (Dougherty and Kawaf 2001).

Experimental method

A section of vertebral bone was imaged using a GE Advantx RFX with a nominal focal spot size of 0.6mm, no intensifying screen and a film image receptor. The focal spot to film distance for this experiment was 200cm.

The bone slice was positioned 50cm from the focal spot, so as to give a geometric magnification of 4× on the film. The radiographic settings used were 50kVp and 800mAs and were chosen to give film densities lying on the linear portion of the film characteristic curve. The film was digitised at 150 dots-per-inch and 8-bits per pixel to give the final image.

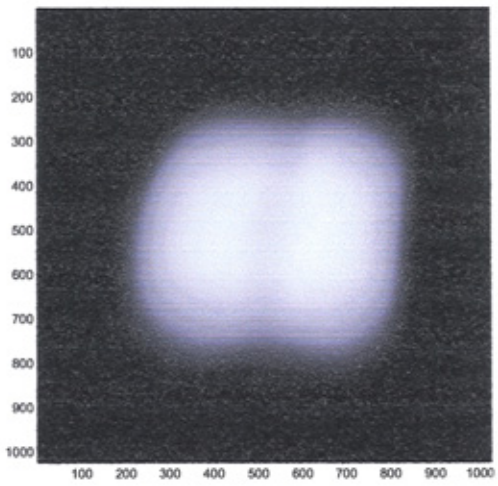

Figure 6.1: Magnified pinhole image of the focal spot showing the bimodal intensity distribution caused by the helical nature of the cathode.

The PSF was obtained by pinhole projection. The pinhole was carefully aligned with the centre of the x-ray beam and 7cm from the focal spot. This configuration gave a geometric magnification of the focal spot of 29× on the film. The magnified pinhole image of the focal spot is shown in figure 6.1. The radiographic settings used to image the focal spot were 50kVp and 1250mAs. The film was digitised at 300 dots-per-inch and 8-bits per pixel. The final dimensions of the bone and PSF images each measured 1024×1024 pixels.

Image processing

The PSF appropriate to the bone image is obtained by rescaling the pinhole image of the focal spot. The rescaling takes into account the difference in the digitisation resolution, the geometric magnification at which the focal spot was imaged and the magnification of the of the focal spot in the plane of the bone slice.

$$\begin{aligned} M_{\text{scale}} &= (\text{Digitisation}) \times (\text{Focal Spot}_{\text{mag}}) \div (\text{Bone Plane}_{\text{mag}}) \\ &= 2 \times 29 \div \left(\frac{150}{50}\right) \\ &= 19.333\ldots \end{aligned} \tag{6.1}$$

The PSF appropriate for the bone image is obtained by reducing the magnified focal spot dimensions by a factor of 19. MEM requires an value for the noise variance in the bone image. This was estimated by analysing the statistics of pixels within small region assumed to have constant signal content.

MEMSYS5 was used to implement a Hybrid maximum entropy deconvolution using the rescaled PSF and noise variance value. The program's Historic MEM switch and setting an ICF (of width 1) was done to implement the Hybrid MEM scheme of section 3.4. The default model was a flat image with low pixel intensities ($m_i = 0.005$).

Results

In figure 6.2 we show the original image and the MEM restoration. The MEM image is smoother in appearance that the original image. To demonstrate the improvement in resolution and visibility of structure in the restored image more effectively we also overlaid some skeletons on each image using Matlab. The skeletons were constructed by determining the mean gray level in each image and then thresholding so that values equal to and above were set to 1, while those below were set to 0. The result is a binary image. The morphological operation of erosion (based on 8-connectedness) was repeatedly applied to this image until a skeleton one pixel wide remained. 'Skeletonisation' removes pixels on the boundaries of objects without allowing objects to break apart. The resulting skeleton is related to the structure of features in the original image. For further discussion of this process see Gonzalez and Woods (2001).

Comparison of the skeletons in figures 6.2b and 6.2d show an obvious difference. The structure in the original image is poorly defined and leads to a skeleton riddled with artefacts and spurious features. In the restored image the skeleton outlines a more regular distribution of the expected intertrabecular spaces which is much more evident than in the original image. The MEM restored image therefore shows more structure and superior detail.

Conclusions

We applied maximum entropy to an image of a vertebral bone slice degraded by substantial geometric blur and demonstrated an improvement in feature visibility and spatial resolution.

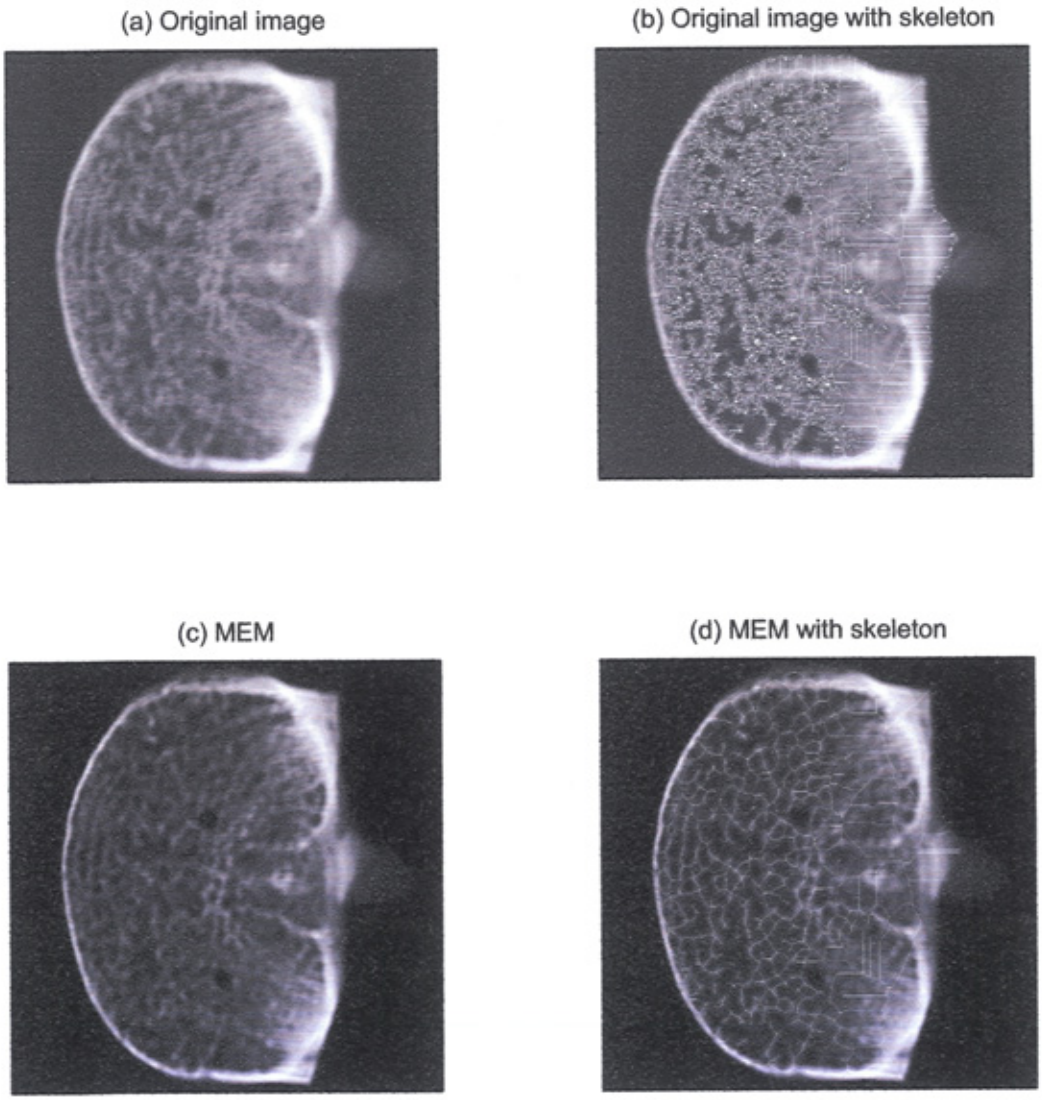

Figure 6.2: (a) is the original image of the magnified bone. (b) is the original image with superimposed skeleton (c) is the result of processing with MEMSYS. (d) is the MEMSYS processed image with a skeleton overlaid on it.

6.2 Mammographic images with simulated microcalcifications

The analysis of MEM in Chapter 4 was applied the TORMAM test object, which contains tiny features and low contrast details against a uniform background. In this section we apply MEM to a selection of digital mammography images with simulated microcalcifications. The mammography images and other data were generously provided by Prof. Dr. ir. H. Bosmans and Dr A.-K. Carton from work on image quality assessment and detector performance in digital mammography (Carton 2004).

Experimental method

The images under consideration in this experiment are a composite of real mammograms with images of simulated microcalcifications whose positions and properties are known with some accuracy. Before proceeding with the application of MEM and subsequent analysis, we describe briefly the methods used to obtain the composite images.

Biopsy specimens of real, malignant microcalcifications were placed on a 4cm PMMA* plate and imaged with an Agfa CR system; the radiographic settings were 27kVp with a Mo/Mo anode-filter combination and geometric magnification 2.0 with the fine focus setting. A high dose (averaging 200μGy) was employed to increase the SNR of the resulting images. Scatter was minimised through use of an air gap rather than a grid. Templates $T_{original}$ containing simulated microcalcifications were constructed using properties such as size, shape and x-ray transmission coefficients of the biopsy specimen images. The microcalcifications in these images suffer from a degree of geometric unsharpness. Idealised templates T_{ideal}, which would be obtained from an ideally sharp image receptor, were calculated by considering the MTF of the Agfa system† and then applying image segmentation routines to isolate the simulated microcalcifications from the background. The T_{ideal} formed the basis from which templates for general use with any other imaging systems (with known MTFs), can be constructed.

The simulated microcalcifications were shown by Carton to be indistinguishable from real microcalcifications. Software phantoms were constructed which consisted of 2cm by 2cm frames, with one or more templates randomly distributed within it. An example image of such a frame is shown in figure 6.4a.

*PMMA: PolyMethyl MetAcrylate

†The actual procedure is similar to the pseudoinverse filter discussed on page 48

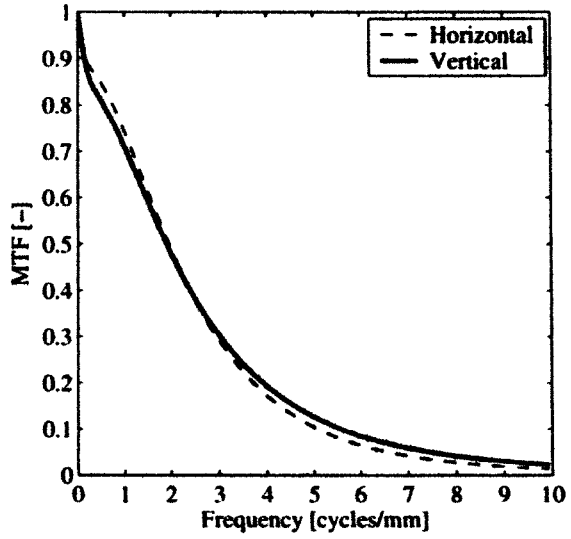

Figure 6.3: The MTF of the Fuji FCR 5000 MA system was measured under conditions as close as possible to those used in clinical practice for the acquisition of a standard breast (Carton 2004). Horizontal and vertical profiles are shown.

The data set for our experiment were mammograms obtained from patients without known pathology. The mammograms were acquired with a Siemens Mammomat 3000 unit under routine clinical conditions: breast compression was used and the exposure was determined by the AEC. The geometric magnification is approximately 1; this was contact *not* magnification mammography (see page 7.1.2). The image receptor was a Fuji HR-BD CR plate, which was readout with a Fuji FCR 5000 MA CR reader. The dimensions of the plate were 18×24cm, giving an image size of 3540×4740 pixels (a pixel size of 50μm).

The software phantoms, based on modifications to T_{ideal}, were rescaled to match the resolution of the mammographic data set and blurred with the MTF for the Fuji FCR 5000 MA system (shown in figure 6.3). The final composite images, which are the basis of our experiment, were obtained by multiplying (element by element) the software phantom image with the raw mammographic images. Each frame measured 200 × 200 pixels in the final image.

The purpose of our experiment is to investigate the effect of maximum entropy deconvolution on the visibility of microcalcifications embedded in a real mammogram. The template type was not varied within a particular

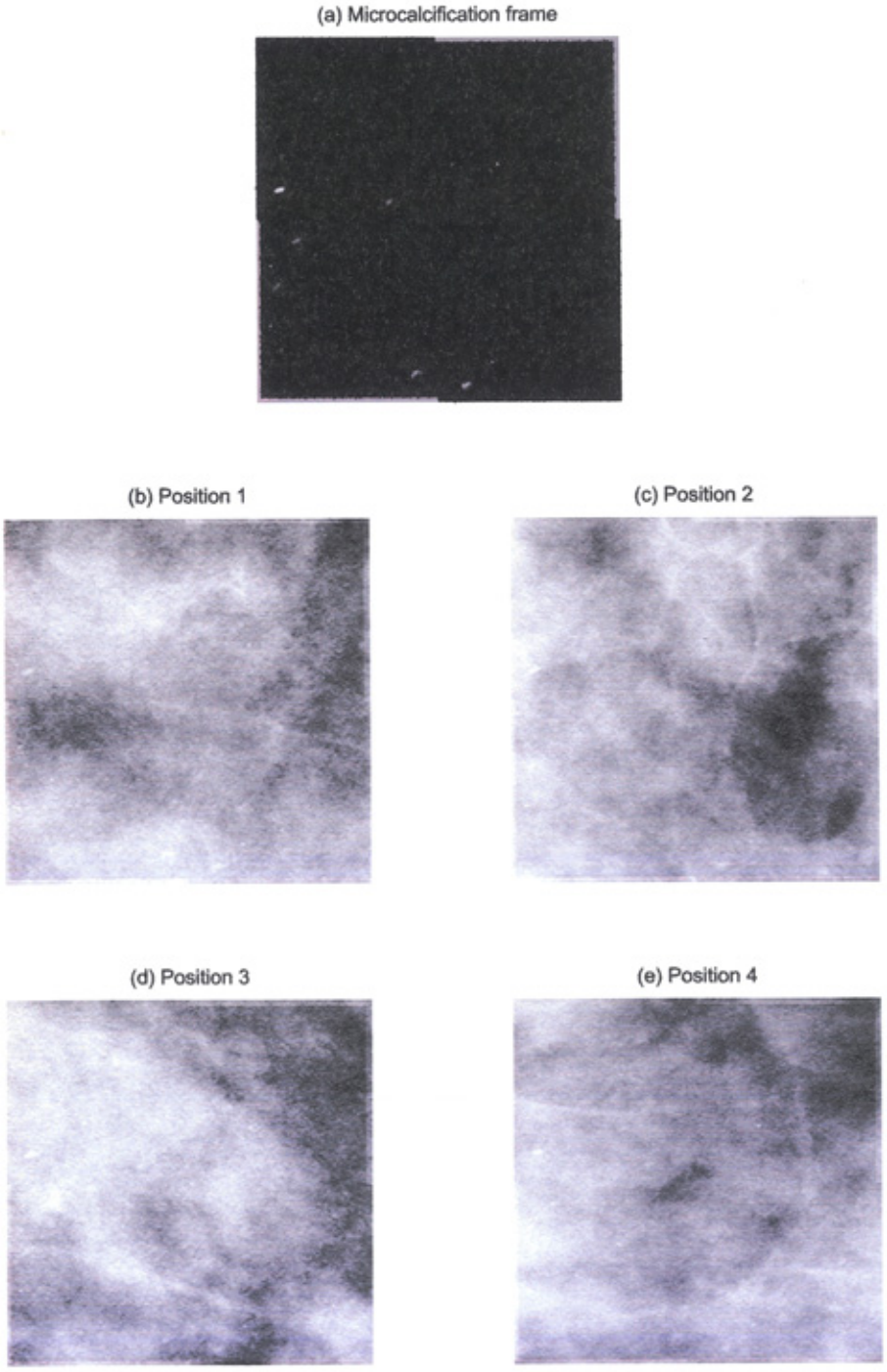

Figure 6.4: An example of mammographic data containing simulated microcalcifications. (a) is an image of the original frame showing the locations of the microcalcifications. The frame was embedded at four locations in the same breast: (b), (c), (d) and (e).

mammogram; so the pattern of microcalcifications was repeated at different locations. Three mammograms were used in the experiment, allowing the placement of the templates against different background types.

1. Almost homogenous fatty tissue.
2. A mixture of dense and fatty tissue without scattered structures.
3. A mixture of dense and fatty tissue with scattered fibroglandular densities.

Image processing procedure

The composite images of mammograms with microcalcifications were in RAW file format; the images were 14-bit grayscale images. The images were converted to DICOM using the Osiris image viewer to make the images compatible with Matlab and MEMSYS5. The approximately linear nature of the image data in previous TORMAM experiments is not applicable in this case as the dynamic range of pixels in these images is much greater. Consequently, the composite images had to be linearised prior to processing with MEM. The conversion to linear pixel values was achieved with:

$$I_{\text{lin}} = 10^{((I_{\log}-8191+\frac{16383}{L}\cdot s_k)\cdot\frac{L}{16383})} \tag{6.2}$$

$$\text{with} \quad s_k = 4-\log\left(\frac{s}{4}\right) \tag{6.3}$$

where L is the latitude and s is the sensitivity of the output data (refer to page 14). Typical values for these images were $L = 2.0$ and $s_k = 0.8$.

A system PSF was constructed directly from figure 6.3; a direct measurement of the focal spot or system PSF was not available to us. This was achieved by measuring points on each curve in figure 6.3 and fitting a polynomial (in a least-squares sense) to these data points. Values between the vertical and horizontal directions were obtained by interpolation. We were able to reconstruct the approximate MTF of the Fuji FCR 5000 MA system. The PSF required by our maximum entropy deconvolution scheme was obtained by taking the inverse Fourier transform of the MTF using Matlab.

For the MEMSYS5 processing we chose a flat default image model (with $m = 0.001$). We used a noise estimate of $\sigma^2 = 0.0001$ obtained by averaging noise estimates within the regions of interest containing the microcalcification templates; the experiment was arranged so that the templates were placed in tissue regions where the system noise would be approximately constant. The full-field images were processed with MEMSYS5 on an Intel Xeon 3.2

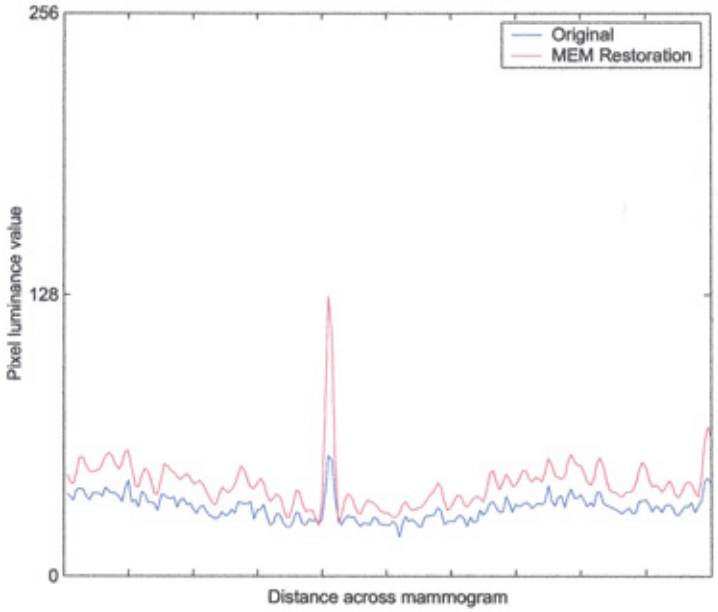

Figure 6.5: Pixel intensity profile through a simulated microcalcification before and after MEM processing.

GHz with 1024MB of RAM. MEMSYS5 converged to the appropriate value of χ^2 in around 9 minutes.

Results

The full-field images were processed but we will show only the cropped regions of interest here.

Figure 6.6 shows regions within Breast A containing microcalcification frames, and the MEM restorations of those regions. Also shown is the original frame without the tissue background. The frame used in conjunction with this mammogram contained 6 microcalcifications; the largest measures 636μm and the smallest is 366μm. The frame was located at two positions in the original mammogram. In the first region (6.6b) we have mainly homogenous fatty tissue. In the second region (6.6d) we have a mixture of fatty and dense tissues.

Figure 6.7 shows regions within Breast B containing microcalcification frames, and the MEM restorations of those regions. Also shown is the original frame without the tissue background. The frame used in conjunction with this mammogram contained 10 microcalcifications; the largest measures

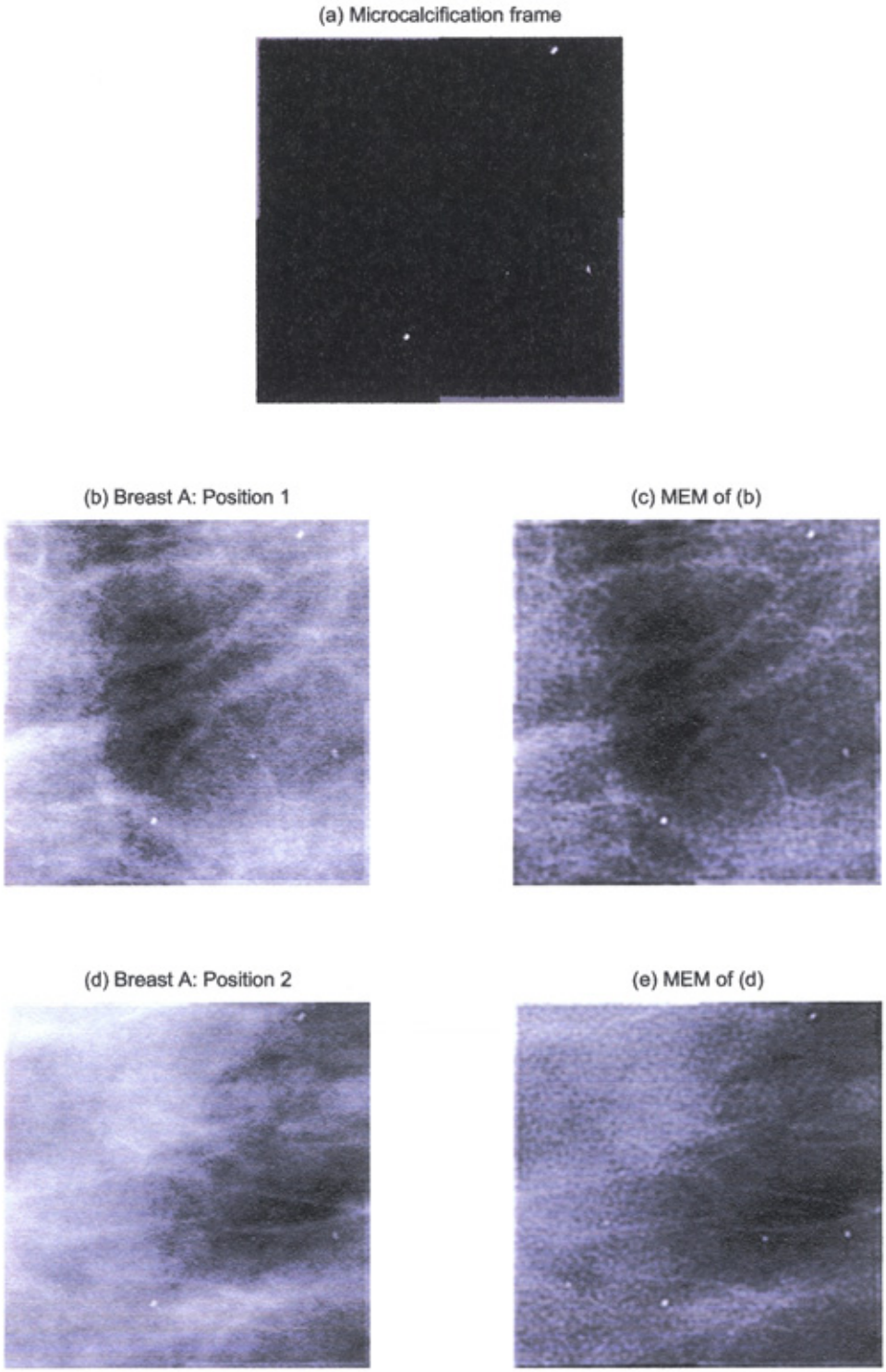

Figure 6.6: Simulated microcalcifications in Breast A. The original frame (a) contains 6 microcalcifications. In (b) and (d) we see the frame embedded at different locations and background tissues. The corresponding MEM restorations are displayed in (c) and (e)

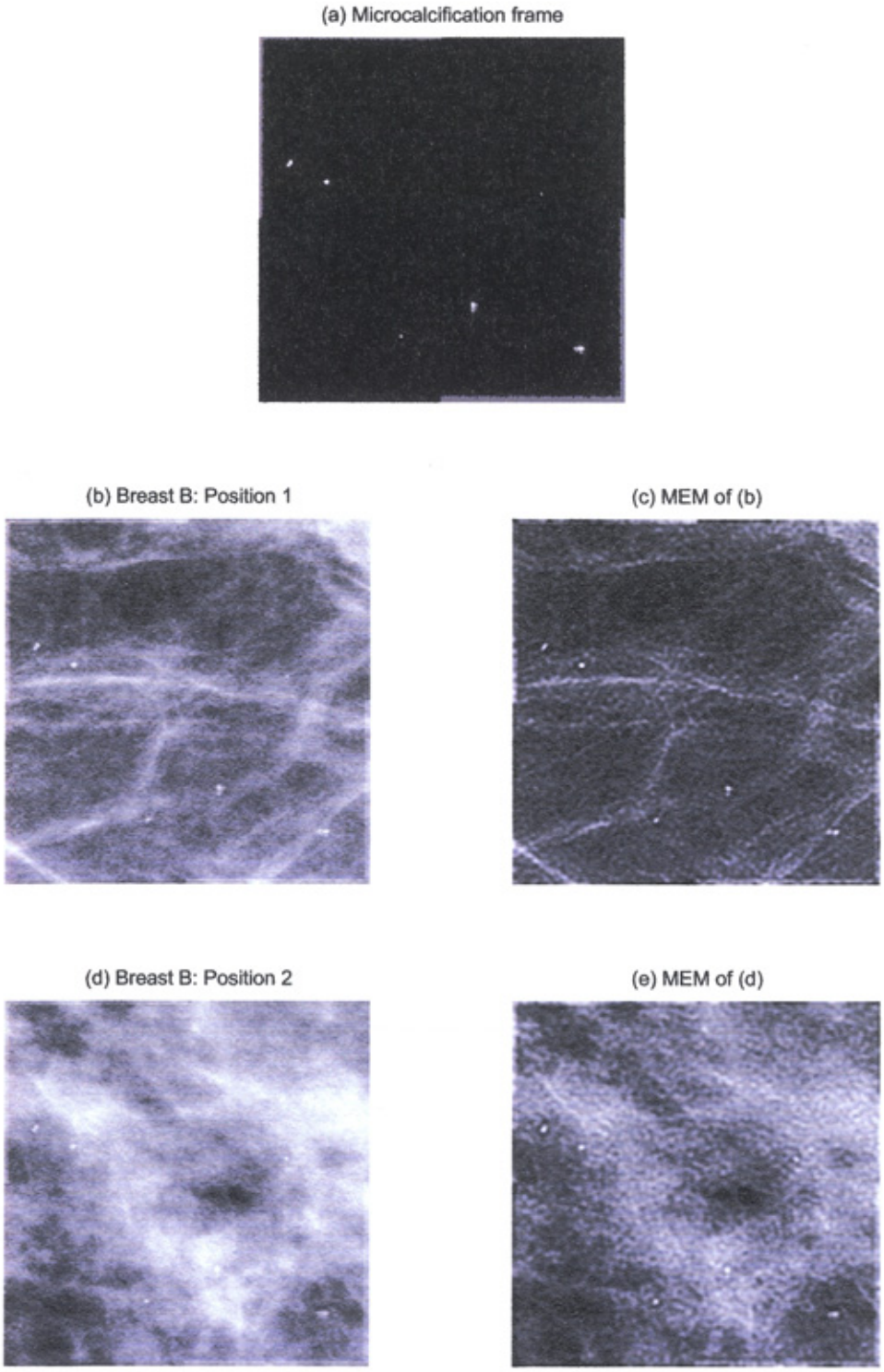

Figure 6.7: Simulated microcalcifications in Breast B. The original frame (a) contains 10 microcalcifications. In (b) and (d) we see the frame embedded at different locations and background tissues. The corresponding MEM restorations are displayed in (c) and (e)

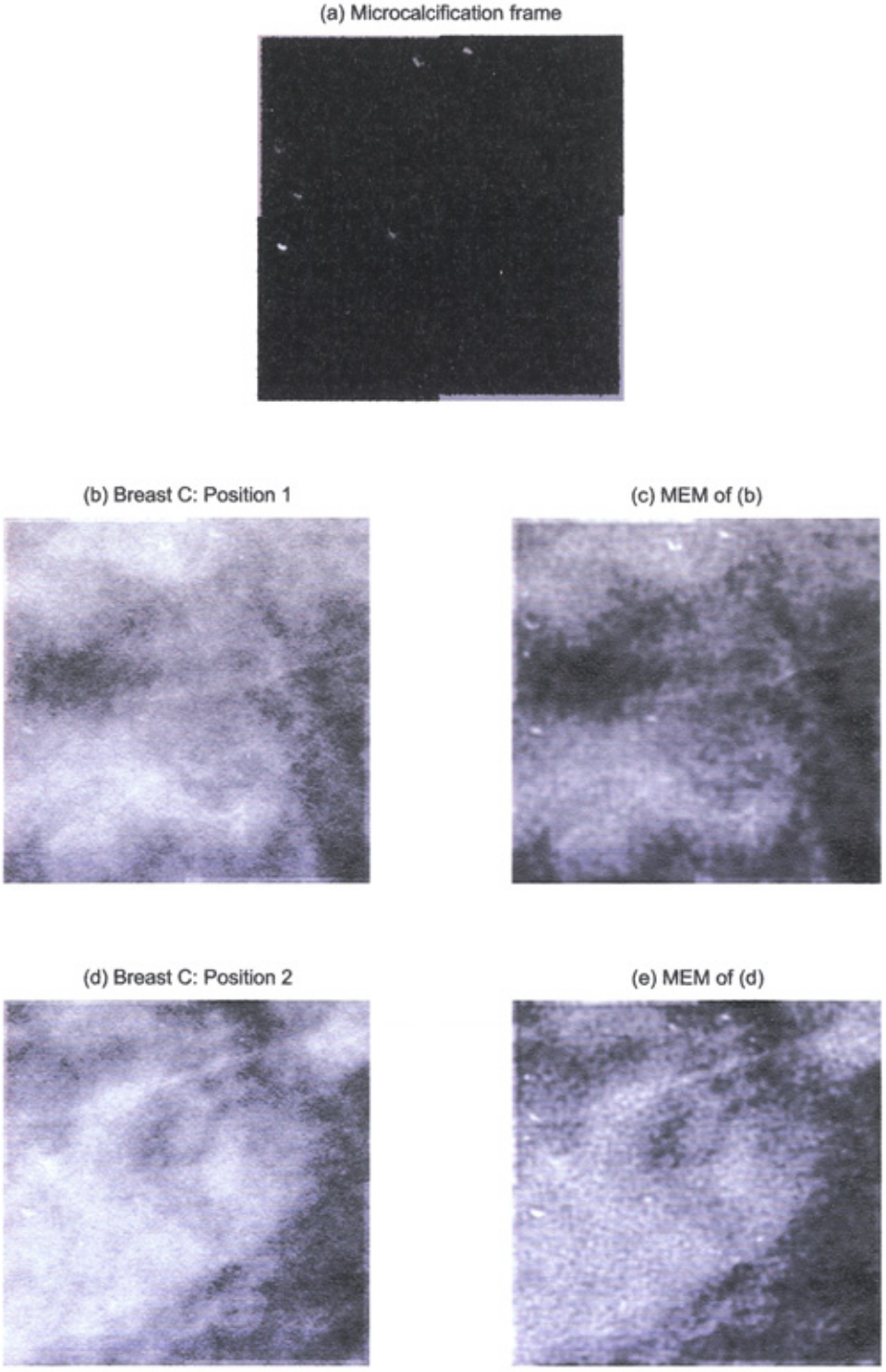

Figure 6.8: Simulated microcalcifications in Breast C. The original frame (a) contains 10 microcalcifications. In (b) and (d) we see the frame embedded at different locations and background tissues. The corresponding MEM restorations are displayed in (c) and (e)

727μm and the smallest is 314μm. The frame was located at four positions in the original mammogram and we show two of them in the figure. In the first region (6.7b) we have mainly fatty tissue with some scattered structures. In the second region (6.6d) we have a mixture of dense and fatty tissue with scattered fibroglandular densities.

Figure 6.8 shows regions within Breast C containing microcalcification frames, and the MEM restorations of those regions. Also shown is the original frame without the tissue background. The frame used in conjunction with this mammogram contained 10 microcalcifications; the largest measures 790μm and the smallest is 357μm. The frame was located at five positions in the original mammogram and we show two of them in the figure. Both regions contain mostly dense tissue with scattered fibroglandular densities.

Conclusions

We applied maximum entropy deconvolution to a range of mammographic data to examine the performance of restoration of images with more realistic properties. Although there is much less blurring in these contact mammography images than those obtained in previous experiments with magnification mammography, we can address image receptor blur and noise reduction using MEM. The improvement in microcalcification signal over the surrounding tissue background is significant (for example, see figure 6.5).

MEM does best on those images where the contrast between microcalcifications and tissue background is already fairly good, i.e. in regions containing mostly homogenous fatty tissues. Improvements in visibility are evident with more complicated backgrounds: where there is a mixture of fatty and dense tissues around the microcalcifications. In regions dominated by extremely dense tissues, MEM restorations do not perform as well as with the other backgrounds.

Some of the restored images of the breast tissues show a somewhat mottled appearance. This is a consequence of deconvolution of structures imaged at the resolution of the image receptor.

6.3 Removal of tomographic blur from a skull phantom

Experimental method

An anthropomorphic skull phantom was placed on top of a table and imaged using a Philips Bucky Diagnost system with linear tomography option. The focal spot to image receptor distance was 100cm and the table-top was 6.5cm above the image receptor. Cut heights through the skull (measured from the table-top) of 2cm, 5cm, 8cm, 11cm, 14cm, 17cm and 20cm were obtained (i.e. 3cm between each plane). The radiographic factors used were 70kVp, 40mAs. These factors were chosen to give realistic levels of quantum noise in the resulting images. The image receptor was a Philips ACR-3 computed radiography system utilising a cassette with dimensions 24cm×30cm and a resolution $R = 65.7$ pixels/cm. The images of the seven planes through these cut heights are shown in figure 6.9.

Blurring functions

The blurring functions are calculated for each of the seven planes using equation 5.7 on page 136. The results are shown in Table 6.1.

	Plane 1	Plane 2	Plane 3	Plane 4	Plane 5	Plane 6	Plane 7
Plane 1	1	163	337	523	723	939	1172
Plane 2	262	1	169	349	542	751	977
Plane 3	419	163	1	175	362	564	782
Plane 4	576	325	169	1	181	376	587
Plane 5	733	487	337	175	1	189	391
Plane 6	890	649	504	349	181	1	196
Plane 7	1046	812	672	523	362	189	1

Table 6.1: The lengths of the blurring functions, measured in pixels, in each plane of the skull phantom. The object in the focal plane is unblurred. This is represented as a convolution with a unit impulse.

Image processing procedure

The skull images were pre-processed using Matlab. The image data were 'cleaned' to remove spurious bright pixels (see page 98) and linearised via equation 4.4 as required by our MEM algorithm.

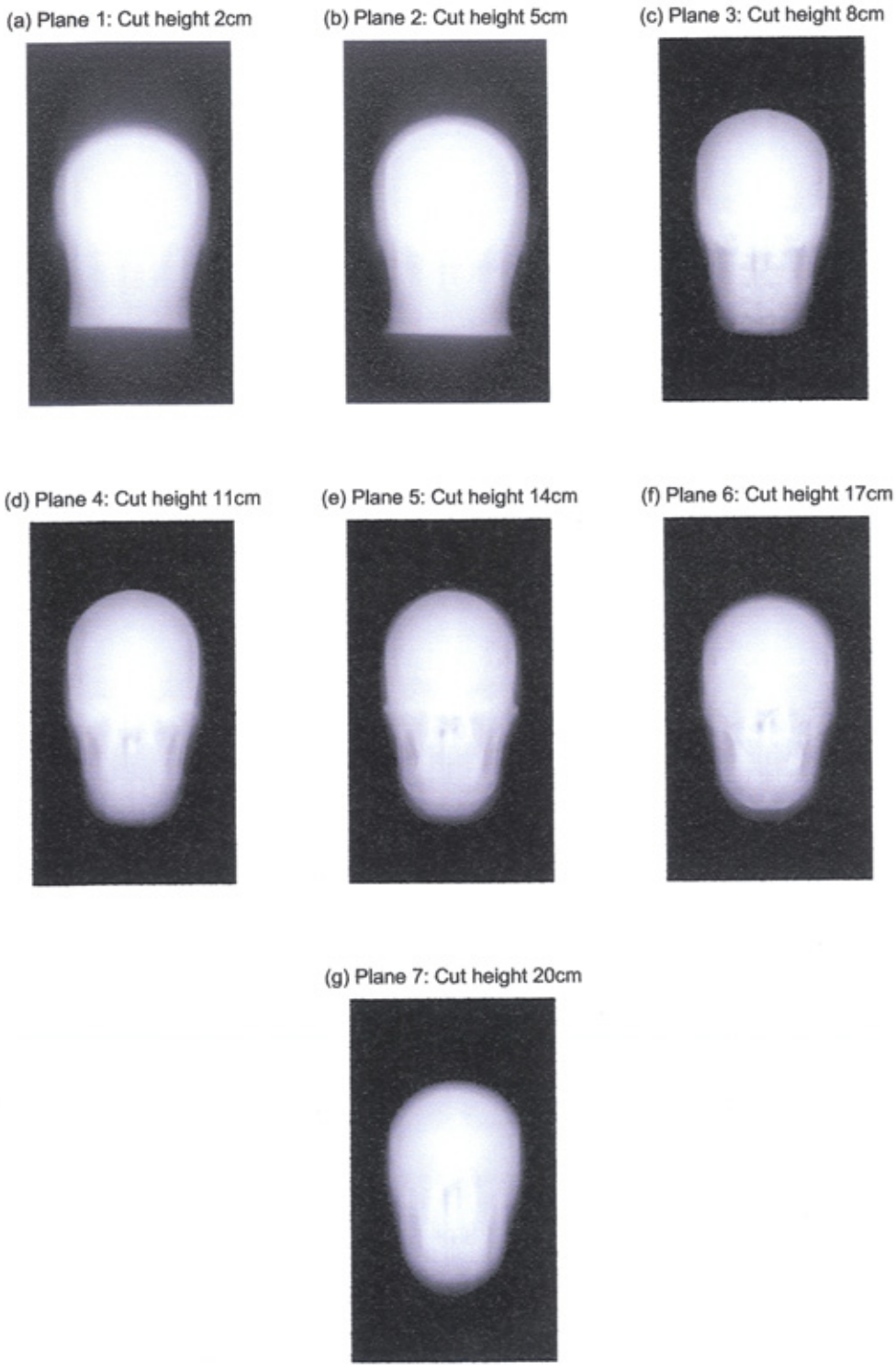

Figure 6.9: The complete set of tomograms obtained for 7 cut-heights through the skull phantom.

MEM requires an estimate of the noise variance which was assumed constant across the mammographic images of chapter 4. This was a reasonable assumption in those situations because the regions classed as 'signal' were not markedly different from regions recognised as being background, in terms of observed pixel luminence values. The situation is different with regard to the skull images considered in this chapter. Regions outside the skull correspond to areas in which the x-ray beam is largely unattenuated (large numbers of detected photons). The estimate of the noise variance, which varies with the mean number of photons, will be somewhat higher in these areas than attenuated within the skull. If we estimate the noise variance based on small regions outside the skull then the effect on the restoration process is rather conservative; more weight is given to the entropy term which ensures that evidence for real structure must be stronger than if the variance estimate was taken from within the skull. While this simple approach is attractive in it's pragmatism - selecting an area of constant signal among the structure of the skull is problematic - it is possible to modify the MEM technique to handle noise levels which vary across the image. A modification of equation 3.11 for this situation is:

$$\chi^2 = \sum_i \frac{(d_i - \hat{d}_i)^2}{\sigma_i^2} \tag{6.4}$$

After linearising the image, pixel values are linearly related to the number of photons (N) detected at the corresponding position on the receptor. The noise variance is proportional to $\sqrt{N}$ (page 16) and so we can estimate expected noise levels for regions within the skull based on their pixel values. A directly measured estimate from a uniform region outside the skull is used to convert the relationship between the relative noise differences to absolute pixel values. Noise variances for the normalised images typically fell in the range 0.0001 to 0.0005. The default image model used was that of $m_i = 0.001$ for all i (a low valued, flat image).

Maximum entropy deconvolution provides a way of removing the out-of-plane blurring from a set of tomographic images. We introduced the HFE filter in chapter 2 and demonstrated an application to tomography on page 133. In this experiment we will compare the performance of HFE filtering with MEM restoration and also assess the usefulness of HFE as a post-MEM processing step. In applying the high-frequency emphasis filter we chose the parameters $a = 0.5$ and $b = 2.0$ (a subjective choice - as is often the case with the enhancement techniques of chapter 2). A second order Butterworth highpass filter (with radius 5% of the image width) was used to construct the HFE filter. A Butterworth filter allows more control over the final HFE filter

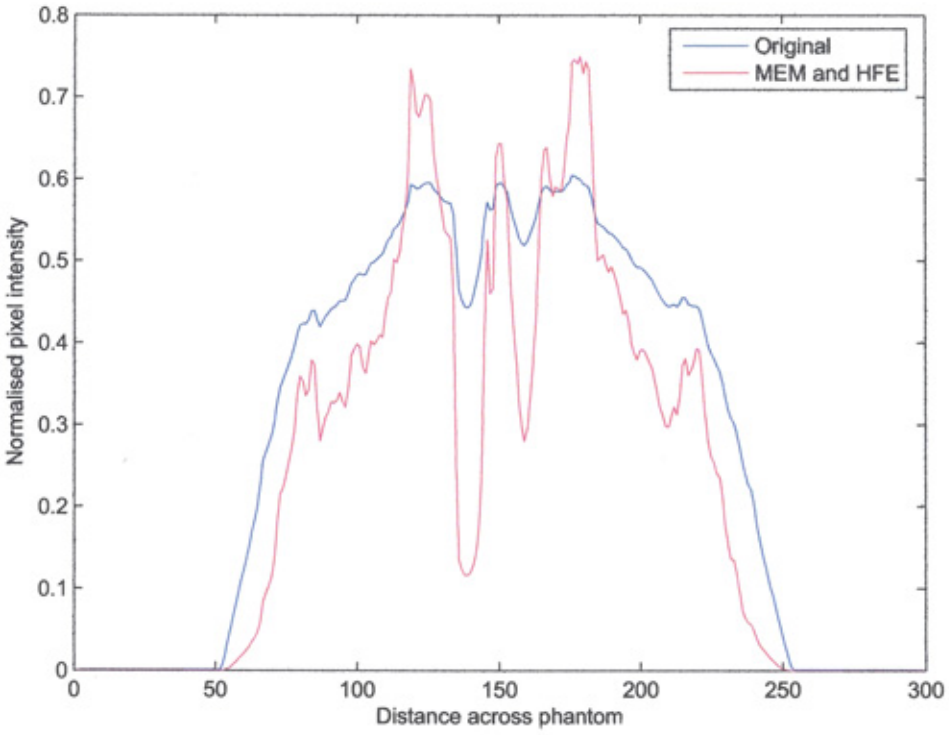

Figure 6.10: Pixel intensity profile through the original and the processed (MEM and HFE) images from plane 5.

'shape' and prevents the introduction of undesirable ringing artefacts. The relative merits of Butterworth, ideal and Gaussian filters in this context is discussed in considerable detail by Gonzalez and Woods (2001). The Matlab code for this 7-plane reconstruction is given in Appendix D.7.

Results

In figures 6.11 to 6.17 we compare the original image, a HFE filtered image of the original, a MEM reconstruction and a post-MEM processed image. The HFE filtered images allow easier inspection of high frequency features in the original image. The MEM images show a substantial reduction in out-of-plane blurring and increased visibility of other features in the skull phantom. The HFE post-processing of the MEM image emphasises the detail high frequency content restored by the MEM restoration.

Examination of pixel profiles across the image in figure 6.10 verify that structures deemed by the algorithm to be outside the focal plane have been reduced in intensity whereas other structures have an increased signal. The overall effect is of increased contrast between out-of-plane blurring and structure in the focal plane. We briefly discuss some of these results in more detail.

With reference to figures 6.11 and 6.12 (Planes 1 and 2). These tomograms show planes 2cm and 5cm above the table respectively, and cut through the back of the skull and top of the spinal column. The original images (a) and (c) can be manipulated[‡] in order to better 'window' the contrast. Even so, details are very difficult to pick out of the overall blurring. The restored images (b) and (d) show spinal vertebrae fairly clearly - this was not easily seen in the original images. Also clearly visible (especially in Plane 2) is the circular outline of the occipital and parietal bones at the back of the skull.

With reference to figures 6.13 and 6.14 (Planes 3 and 4). These tomograms show planes 8cm and 11cm above the table respectively. The neck vertebrae are visible in the original image (a) from Plane 3 but are much more clearly defined in the MEM restoration (b). The nasal bone and cavity is beginning to emerge from the background blur in Plane 4. The sphenoid bone, which contributes to each orbit is clearly seen in (d).

With reference to figures 6.15 and 6.16 (Planes 5 and 6). These tomograms show planes 14cm and 17cm above the table respectively. The mandible becomes more prominent as we move upwards through the skull towards the face-side. Removal of blur allows a better view of the nasal cavity (showing two prominent holes) and some of the teeth. The bones which contribute to the circular orbits are visible in the original plane 6 tomogram but they become much clearer after processing.

With reference to figure 6.17 (Plane 7). The final image shows a plane 20cm above the table. This plane cuts through the skull just below the face. The nasal cavity is again prominent but the orbits, which are mostly below the cut, are not. The reduction in out-of-plane blurring enables the teeth to be more clearly visible.

Discussion

We approximated the true 3-D object distribution of a skull phantom with seven equally spaced 2-D planes 3cm apart. Despite the crudeness of this approximation our deconvolution technique improves the contrast of features within each focal plane while suppressing blur outside it (see the pixel profiles given in figure 6.14c for evidence of this). We compared coronal MRI scans[§]

[‡]The Osiris image viewer was used for this purpose.

[§]Available online from University of Kansas Medical Centre: `http://www.kumc.edu/`

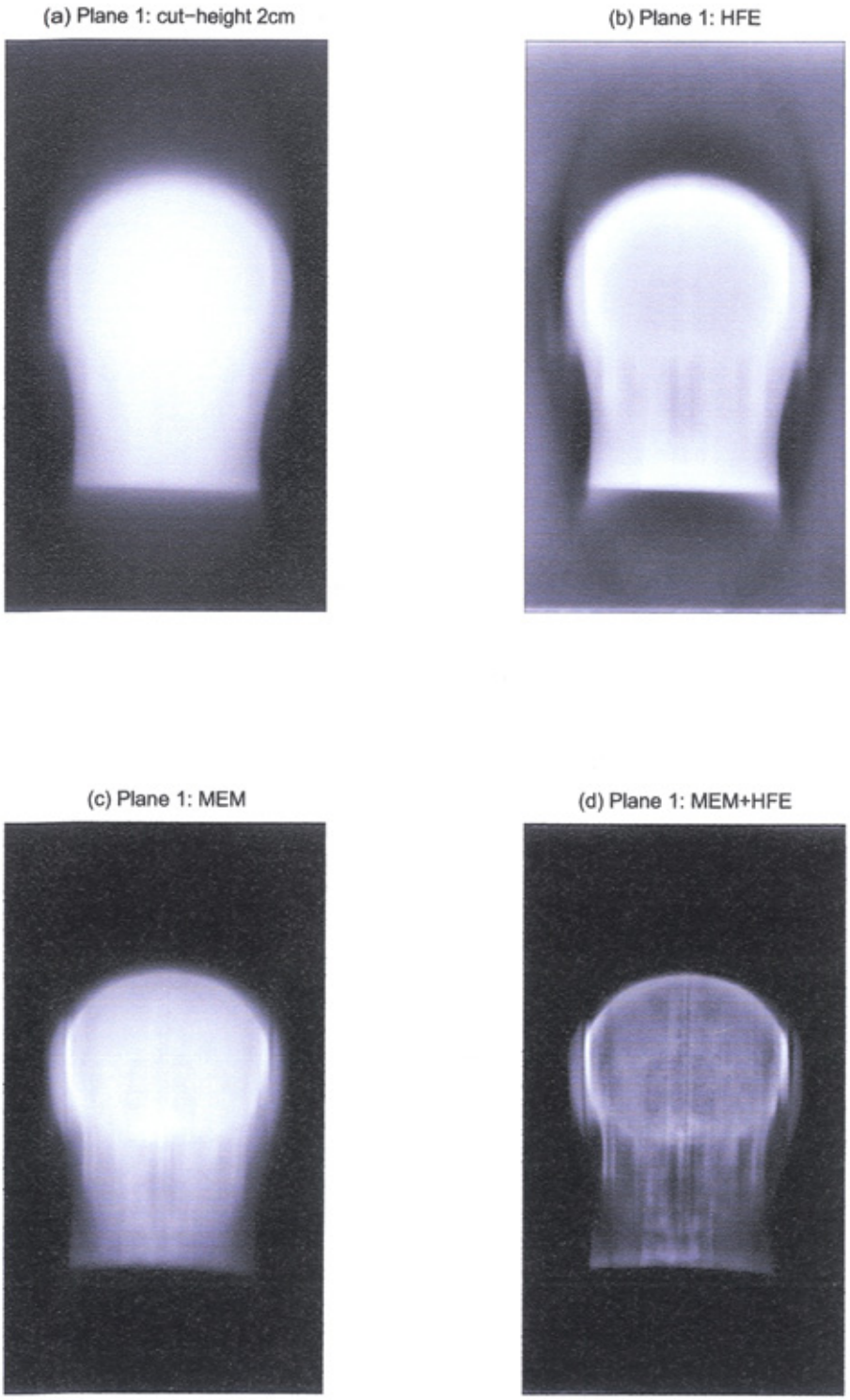

Figure 6.11: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

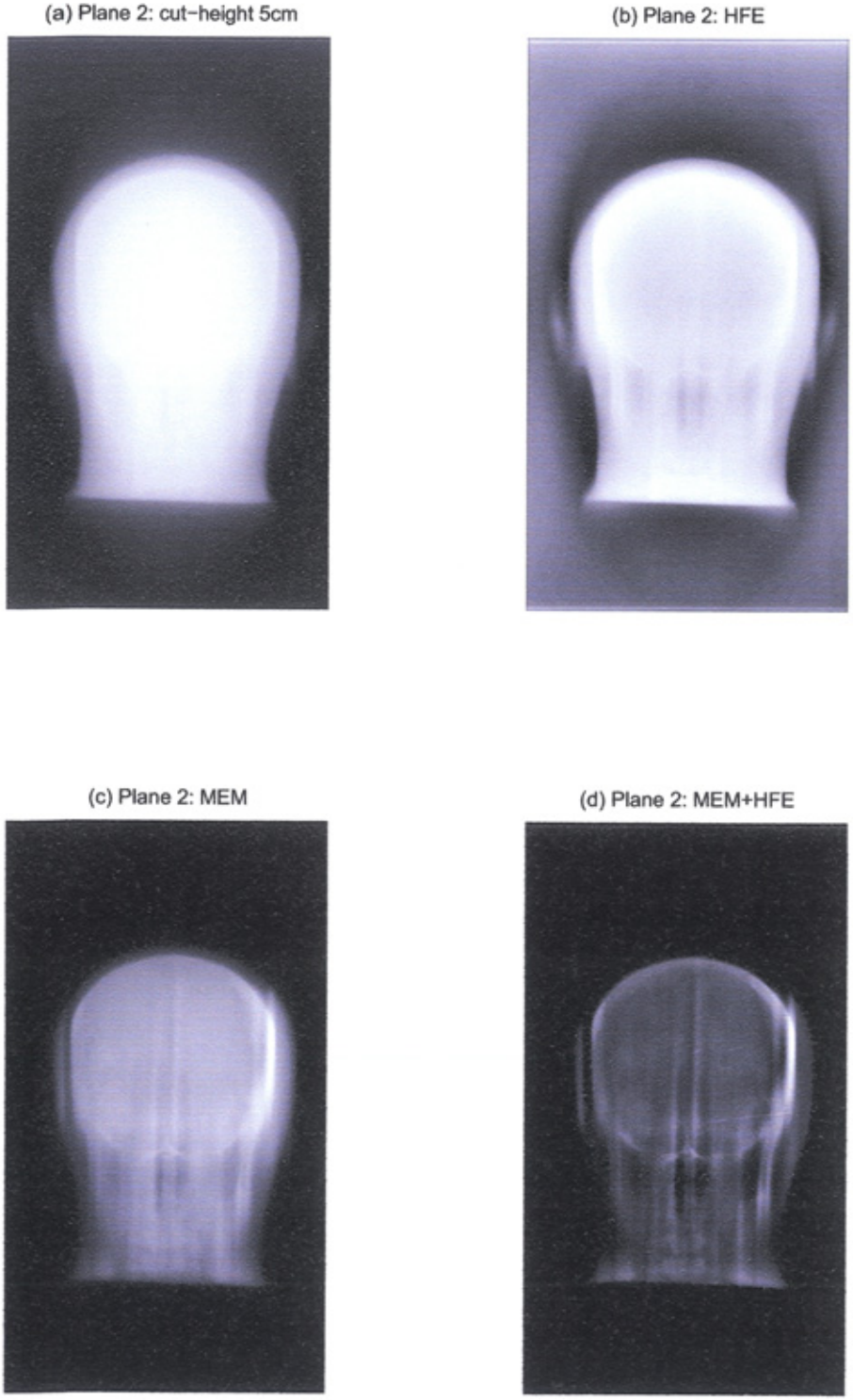

Figure 6.12: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

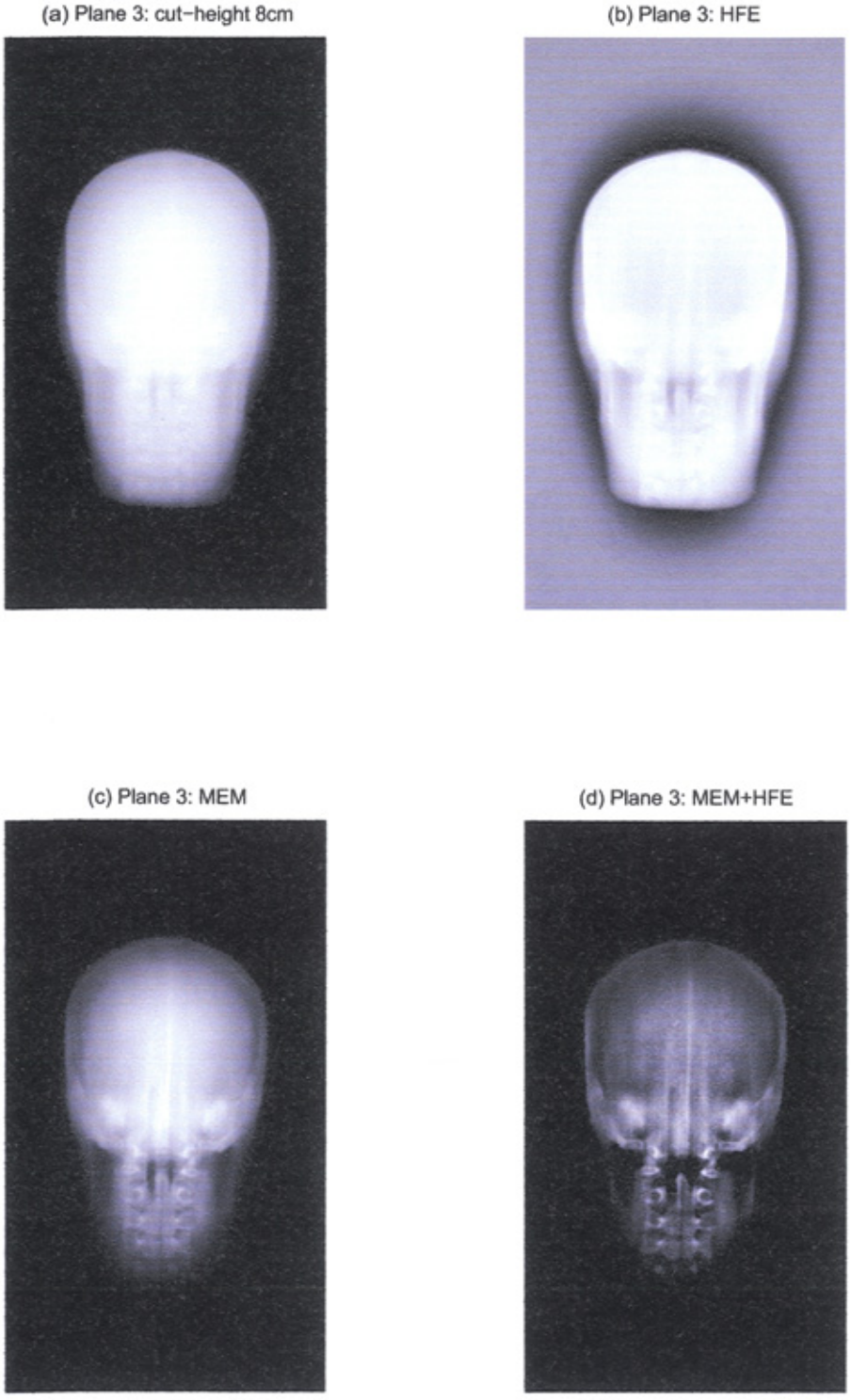

Figure 6.13: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

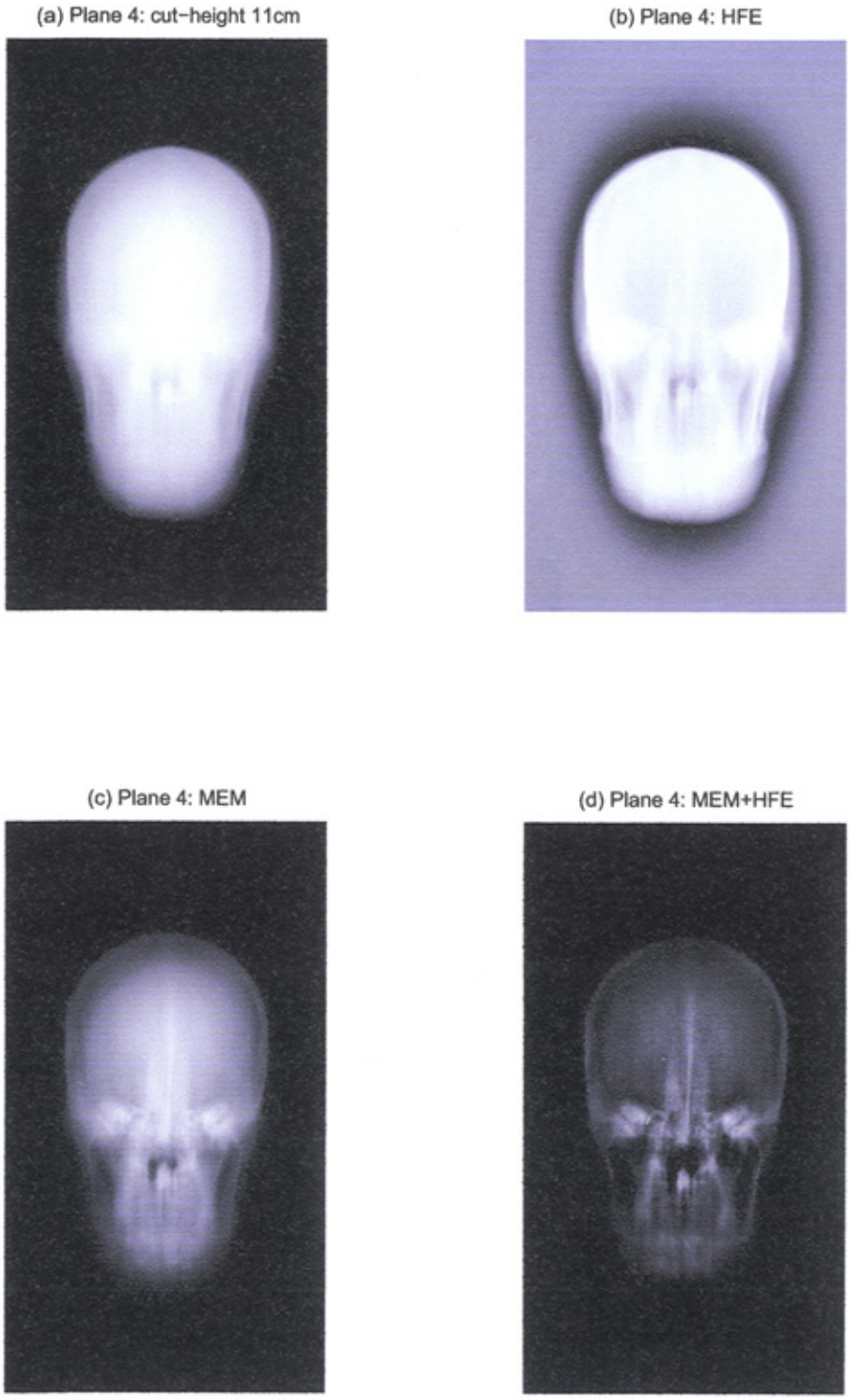

Figure 6.14: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

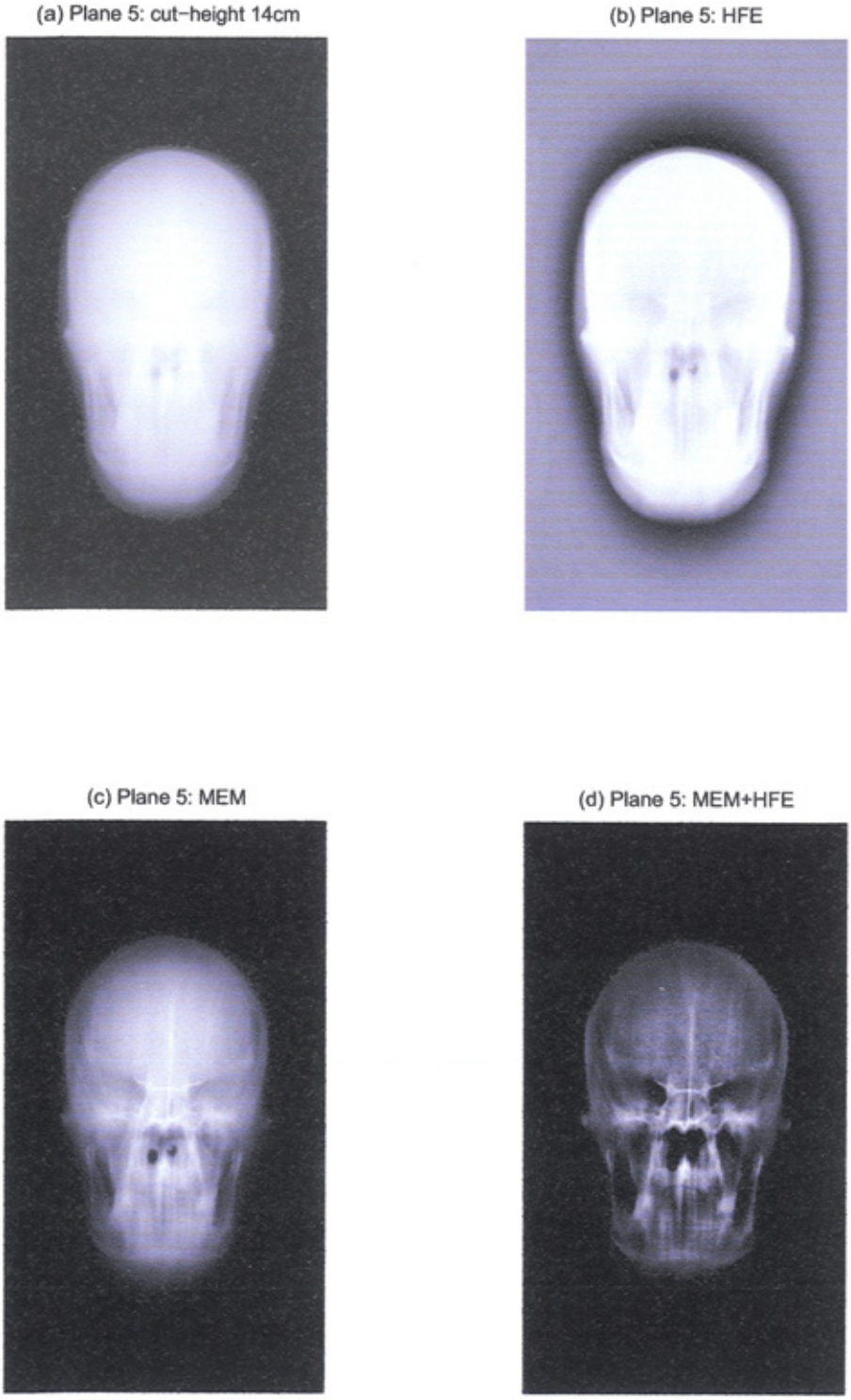

Figure 6.15: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

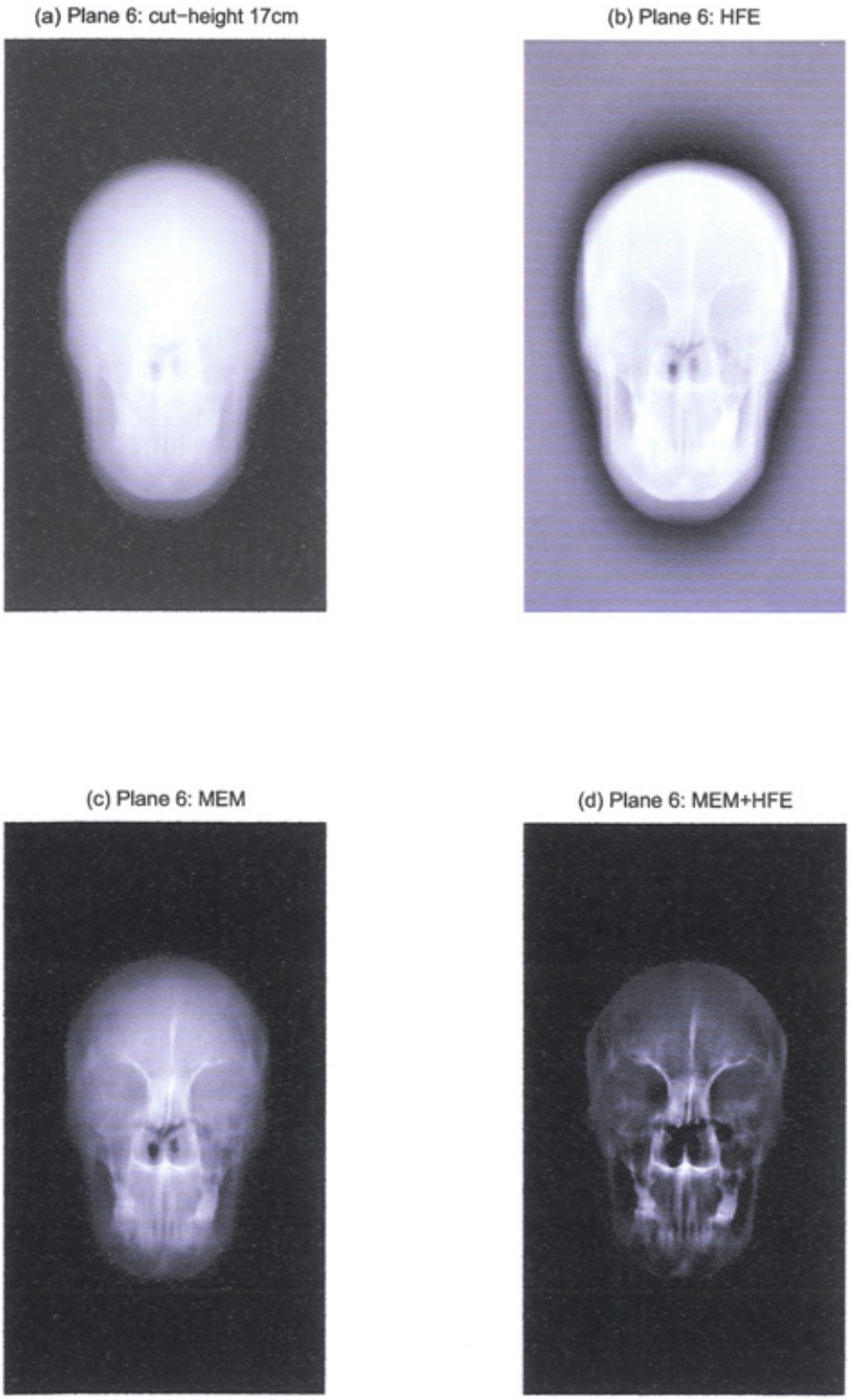

Figure 6.16: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

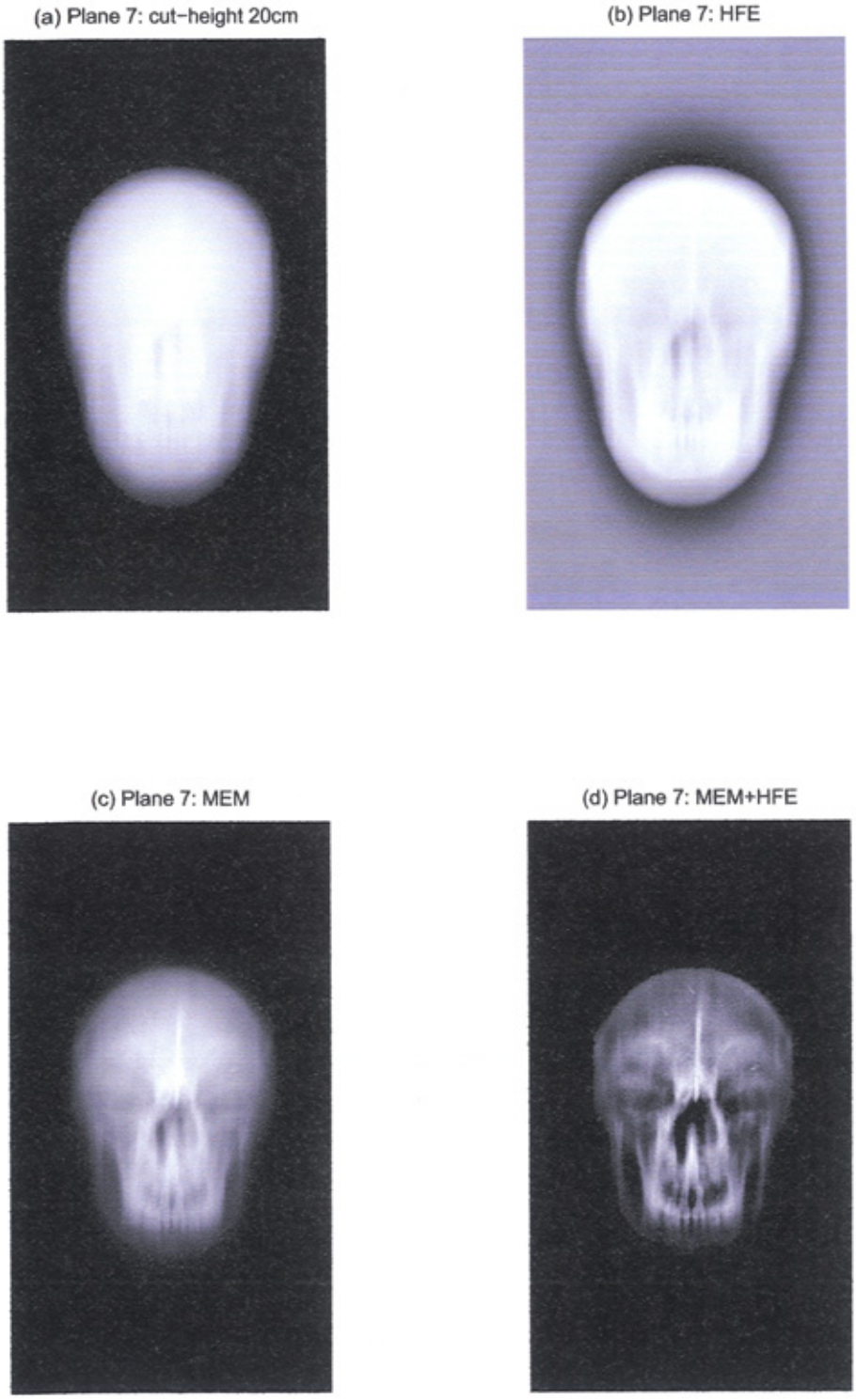

Figure 6.17: The original tomogram of the skull phantom is shown in (a). The result of HFE filtering the original image is shown in (b). The application of our MEM scheme to the original image resulted in the image (c). The image resulting from post-processing the MEM image with HFE is shown in (d).

of a patient's head with our tomographic images to assess the visibility of features in the focal plane. This is admittedly an unsophisticated comparison as the skull phantom does not contain a brain. Nevertheless, it gives a good indication of which bones lie within the focal plane. Despite the advantages of MEM processing some weaknesses of the linear tomographic method remain an issue. For example, the presence of striping artefacts associated with linear motion are largely unaffected by the restoration process.

Our noise model in these experiments, which takes into account the varying noise levels within the image, differs from the one used in chapter 4. However, in comparing the results shown in figures 6.11 to 6.17 with test reconstructions based on a forward map assuming a constant noise variance we can see no discernible differences in the restorations.

CHAPTER 7

Discussion and Conclusions

In this work we have processed images, of both artificial and clinical origin, from medical imaging techniques in which maximum entropy deconvolution (and our Hybrid MEM scheme in particular) would be of immediate benefit. In this final chapter we summarise our findings and discuss possibilities for further research in these and related areas.

7.1 Magnification mammography

7.1.1 System optimisation options

The basic dilemma in magnification mammography is the trade-off between exposure time and focal spot size. The long exposure times associated with small focal spots can be self-defeating, in that patient movement can then become more problematic than geometric blurring. The magnification mammography technique has therefore become restricted because of constraints on the radiographic settings (focal spot size, magnification and exposure). A working MEM deconvolution scheme allows new possibilities for optimising the radiographic settings. In our experiments we investigated how MEM processing might be used to offset any detrimental effects to the resulting image when one or more of the standard radiographic settings were changed. We summarise these findings below.

Focal spot size

Our MEM based approach has been applied to the problem of removing the geometric blurring associated with the finite size of the focal spot in magnification mammography. The TORMAM results of section 4.2.1 show an improvement in image spatial resolution and an improvement in terms of the image signal-to-noise ratio perceived by the observer, as evaluated using a standard phantom and at a realistic quantum noise level. Measurements of the improvements support our claim that images obtained at conventional 1.8 magnification, but using a broad focal spot, can be processed to be as good as images that would be obtained with a fine focal spot.

This successful demonstration of image de-blurring in noisy radiological images offers the possibility of weakening the link between focal spot size and geometric blurring in radiology, and thus opening up new approaches to system optimisation.

Dose reduction

We applied maximum entropy deconvolution to a series of images obtained with successively reduced x-ray doses (and increasing noise levels). Our results showed that MEM could compensate for the reduced image quality: we quantified this improvement by scoring the visibility of fine details and low contrast features of the TORMAM phantom. The improvements in scores arose because the SNR of features in the TORMAM phantom had been improved over the original images. In our experiments we found that the reduction in image quality caused by lowering the dose from 50mAs to 20mAs could be offset by MEM processing. A reduction in dose from 50mAs to 32mAs resulted in higher scores being assigned to features in the MEM processed image. The lowest dose image (10mAs) did not display any improvement through processing.

Increased magnifications

There are obvious consequences for dose reduction in radiography. Obviously, a lower dose of radiation would be beneficial to the health of a patient. Of particular relevance to the magnification mammography work is that a reduced dose - leading to a skin dose reduction - would be necessary for the higher geometric magnifications considered in chapter 4 to become useful in a clinical setting.

Conventional mammography set-up

In the magnification mammography experiment described in section 4.2, MEM processing of TORMAM phantom images obtained at the conventional dose, 1.8 magnification and with a fine focal spot showed an improvement in feature visibility compared to the original image. The improvement was quantified using a clinical test object.

The digital mammograms with simulated microcalcifications (Carton 2004) which were obtained without magnification (in section 7.1.2 also showed an improvement in the visibility of those microcalcifications after MEM deconvolution. We point out that variation of PSF with depth is probably more critical in these images and that future experiments with contact mammograms should take account of that variation.

7.1.2 Further research and refinements

There are several ways to refine the experiments in deconvolution of magnification mammography images described in this work. We briefly outline some possibilities here.

System optimisation options Further exploration of the relationship between focal spot size, dose and magnification in combination with deconvolution. The comparison of broad and fine focal spots seen in Chapter 4 employed the same dose. The consequences of using a broad focal spot with a short exposure and investigating the effectiveness of deconvolution could be a useful experiment.

Implementing high magnification safely At high magnification the intensity of the skin dose would be relatively high. A reduction of x-ray dose might allow use of higher magnifications with the resulting loss of image quality compensated for by MEM image restoration.

Better focal spot definition The PSFs used to deconvolve the images shown in Chapter 4 were obtained by placing a pinhole in the plane of the object (either separately or at the same time). Ideally, the PSF should be obtained by using a magnified pinhole projection method and then rescaling the image according to the geometric set-up. A PSF obtained this way would be less susceptible to degradations from noise and would provide a better measurement of the focal spot contribution to the blurring.

3D varying PSF One of the assumptions we made in discussing restoration methods in Chapters 2 and 4 was of the invariance of the PSF across the

image. However, images obtained with a multi-pinhole setup showed that the PSF did vary somewhat in shape and intensity across the image. A future refinement to this work would take that variation into consideration. The depth-varying nature of the PSF is another factor which might be accounted for as part of some future research. When the breast is compressed in mammography there will be a variation in the amount of blur in the resulting mammogram because some features in the breast were actually closer to the focal spot than others. The rate at which the PSF size varies with depth is negligible in the magnification mammography setup. However, in contact mammography this variation is considerably greater and some allowance should be made in the forward map for the image formation process. For the deconvolution of mammograms in section 6.2 we assumed the dominant source of blur was image receptor blur. Focal spot* blurring would also have been present to some extent and the method might have yielded better images if some account of PSF and the variation of it across the breast thickness had been observed.

Additional processing There are a number of image processing techniques that could be used alongside MEM to enhance the performance of the processing described in this work. High frequency emphasis filters might be used as a post-MEM processing step to enhance the appearance of microcalcifications. Wavelet analysis also allows for the extraction of small-scale features after MEM has been applied. The use of wavelets to denoise an image prior to MEM processing might also be a fruitful approach (see the examples in section 2.2.2). Alternatively, wavelet-based maximum entropy deconvolution (Starck and Murtagh 1994; Starck and Pantin 1996) techniques exist. Applications to other imaging problems have included improvement of maps of the Cosmic Microwave Background (Vielva et al. 2001).

7.2 Linear tomography

We described a technique for modeling the linear tomography imaging system and modified our Hybrid MEM algorithm to be consistent with this model. Tomograms processed with our technique showed a substantial reduction in the blurring associated with structures outside of the focal plane. A further improvement in the visibility of features was seen after post-processing with a high frequency emphasis filter.

*Focal spot measurements, in this case, were not available to us.

7.2.1 Further research

There are several further ideas which could potentially benefit the tomography processing techniques discussed in this work:

Reduced swing angle Experiments to gauge the usefulness of our technique using smaller tomographic swing angles (perhaps $10° - 30°$). Smaller swing angles result in thicker tomographic 'slices'

Focal spot blurring In our model formulation it was assumed that geometric unsharpness was negligible. However, the magnifications encountered in the 7 planes of the skull phantom ranged from $1.1\times$ (nearest the image receptor) to $1.4\times$ in the plane closest to the focal spot. A further refinement to the technique would be to incorporate geometric blurring deconvolution into the formulation (possibly as a post-MEM processing stage).

Other post-processing methods MEM reduces a substantial amount of out-of-plane blurring and we used simple Fourier techniques to enhance the restored images. Post-processing using other techniques (wavelet transforms, for example) might also have been used.

More complex motion Appropriate modification of the forward map would make the technique amenable to more complex types of focal spot/image receptor motion.

It would be interesting to see if the acquisition of fewer planes, which would allow for a reduced dose to the patient, could reconstruct more planes than we have tomographic data for. For example, how well could we reconstruct 7 planes from 3 tomograms? Tomosynthesis techniques (page 131) which allow the generation of any plane retrospectively is a better alternative if available. However, we will discuss potential applications to a similar blurring problem in dental radiography on page 179.

7.3 Deconvolution: further applications

Heavier beam filtration

Although there has been recent research into the production and use of monochromatic x-rays in digital mammography (Lawaczeck et al. 2005) it is usually the case that ordinary x-ray tubes produce a broad energy spectrum (similar to those seen in section 1.2). Filters modify the x-ray spectrum by reducing

the non-image contributing low energy photons (which would be absorbed by the breast) and suppressing the contrast-reducing high energy photons. The normal method of modifying the spectrum is to introduce metallic foils of high atomic number into the beam path. There has been considerable research into the potential savings in patient dose by optimisation of the energy spectrum to suit the particular imaging application (Shrimpton et al. 1988; Sandborg et al. 1989; Hansson et al. 1997). In mammographic imaging the target/filter combination is usually molybdenum or rhodium. Monte Carlo techniques to evaluate the effects of target/filter combinations on factors including average glandular dose were carried out by Dance et al. (2000b). Analysis of mammographic image data to gauge the effect of different target and filter combinations on patient dose were published by Chevalier et al. (2004). The dose reducing effects and improvement to image contrast of aluminium, copper, titanium and gold filters in abdominal and dental radiography was discussed by Sandborg et al. (1994). In other dose reduction investigations in dental radiography, filtering by niobium (Calicchia et al. 1996) and the rare-earth metal erbium (Horner et al. 1988) have proved to be useful.

There is attenuation of energy across the whole spectrum when filtration is used. The total output power, which is crucial to the image formation process, is reduced by filtration. The power can increased to compensate for this, but the resulting heat load on the target can cause problems; there are practical limits to how heavily the beam can be filtered. The use of a broad focal spot would go some way to prevent the possibility of tube burn-out, but as we have seen already, this introduces greater penumbral blurring into the final image. However, the blurring introduced by a broad focal spot can be compensated for without loss of image quality in a mammographic setting by post processing with MEM (Jannetta et al. 2004). The application of maximum entropy deconvolution, in conjunction with the use of broad focal spots, may open up new optimisation research into filter choices allowing heavier filtration of the x-ray beam.

Scatter deconvolution

The causes and effects of radiation scatter were introduced in Chapter 1 (page 9). In the mammography and tomography experiments considered in this work, image degradation due to scatter was not specifically addressed in either of the image formation models for those problems; instead it was treated as an inseparable component of the overall noise which also comprised quantum noise and image receptor noise.

Scatter can be treated as another source of blurring and this approach

has been used to apply deconvolution to images degraded by scatter. A technique similar in approach to the inverse frequency filters discussed in Chapter 2 has previously been applied to images of suitable test objects (Seibert and Boone 1988). The scatter PSF was modeled as a modified Gaussian distribution whose 'shape' was fitted to measurements taken from images of the test object. Seibert's results showed a reduction of scatter and an improvement in radiographic contrast. An improved method using an edge-spread technique to measure the ratio of scattered radiation to primary radiation is discussed by Cooper et al. (2000).

Tomosynthesis and 3D breast imaging

Our linear tomography scheme for reducing the out-of-plane blur could be adapted to treat images obtained with tomosynthesis (described on page 131). There is a degree of overlap with our mammography work as there is a large amount of recent research in applying 3D techniques to breast imaging. In conventional mammography, the overlapping structures within the breast can potentially hide microcalcifications or produce shadows in the resulting images which may mimic a lesion. Tomosynthesis provides a method of overcoming these problems. A description of digital tomosynthesis techniques applied to breast imaging is given by Niklason et al. (1997). See also Simonetti et al. (1998); Chen and Ning (2003); Gong et al. (2004). Removal of tomographic blur from mammograms with our modified MEM scheme could potentially increase feature visibility while at the same time increasing SNR of microcalcifications in the focal plane.

Dental panoramic tomography

Dental panoramic tomography (DPT), also called orthopantomography, is a technique which is widely used to obtain radiographs of the teeth and jaw. We describe the technique briefly here, but more detailed introductions and reviews of DPT can be found elsewhere: for example, refer to Murray and Whyte (2001) and Moore (2002). DPT has traditionally used film-screen image receptors but the merits of digital receptors are currently being explored (Carmichael et al. 2000; Ekestubbe et al. 2003) with some studies concluding that digital systems are as accurate as film systems (Wenzel 2000; Molander et al. 2004; Luangjamekorn et al. 2005). Advances in the methods and technology of digital DPT have been described in recent literature: see Mol (2004) and Hatcher and Aboudara (2004).

The implementation of DPT is similar to the tomography techniques described in Chapter 1. The DPT machine itself is constructed in such a way

that the movement of focal spot and image receptor follows the profile of the mandible. DPT systems map the 'curved' shape of the jaw to a long, 'straight' projection on the receptor. All the teeth, both upper and lower, are effectively in the same image layer and in sharp focus. Features outside the focal layer are blurred and distorted to varying degrees; the effect of out-of-layer blur is to reduce the visibility of the sharp, in-focus features. The blurring associated with DPT is only slightly more complicated than the linear tomography situation seen in Chapter 5. Most of the blurring is obviously due to the bulk of attenuating bone and tissue being 'behind' the teeth. The application of MEM in our work used several image planes to iteratively restore the postulated 'true' slices. The method would not require much modification if additional image layers were available in a DPT situation. This would obviously result in a higher patient dose so is not a particularly feasible approach. In the usual case where a single image is obtained then the situation is very ill-posed. However, the ability to incorporate prior knowledge into the formulation gives MEM an advantage over other deconvolution techniques; perhaps a crude estimate of the attenuating properties of those layers beyond the focal-layer would allow a modified MEM scheme to converge to solutions consistent with the observed image. Some of the image enhancement methods discussed in Chapters 2 and 5, such as the wavelet based removal of blur (page 132) or the high frequency emphasis filter (page 133) might also prove to be useful, either alone or as post-processing steps to MEM.

Digital angiography

Angiography is a technique to image the blood vessels of the body by injecting a contrast medium/dye thus rendering them highly attenuating relative to surrounding tissue and bone. In the original technique, now referred to as digital subtraction angiography, a mask image is subtracted from an image of the same region in which the attenuating dye is flowing moments later. In the resulting image the blood vessels are the most prominent feature. There are several variants of this technique including the use of, for example, MRI technology rather than an x-ray source and now also without image subtraction.

There is a potential application of MEM deconvolution to 3D rotational angiography (Cornelis et al. 1972; Kumazaki 1989, 1991). In this mode of imaging multiple views are obtained following a single injection of the contrast medium. However, an increased radiation dose and prolonged injection of the dye are necessary for useful images to be obtained. Attempts to reduce the inherent problems of noise and artefacts using simple filtering

techniques have been described in recent literature (Meijering et al. 2002; Wong and Chung 2004). The future application of MEM deconvolution to 3D rotational angiography may have similar benefits for system optimisation to those already discussed for magnification mammography.

7.4 Conclusions

In this work we have investigated several medical imaging problems, with particular emphasis on magnification mammography and linear tomography with CR. We have shown that a modified 'hybrid' version of the historic maximum entropy method can provide a radiographer with more flexibility in choosing radiographic settings in magnification mammography. With appropriate modification of the Hybrid method in relation to linear tomography image formation process, a substantial reduction in out-of-plane blurring is possible.

More generally, the demonstration of applicability of MEM in noisy radiographic images suggests that this technique could be a useful tool in the optimisation of digital radiology.

APPENDIX A

The Fourier Transform

The Fourier transform $F(u,v)$ of a two variable, continuous function $f(x,y)$ is defined as:

$$F(u,v) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)e^{-i2\pi(ux+vy)}\,dxdy \tag{A.1}$$

where $i = \sqrt{-1}$. Conversely, $f(x,y)$ can be recovered from a given $F(u,v)$ by application of the inverse transform:

$$f(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} F(u,v)e^{i2\pi(ux+vy)}\,dudv \tag{A.2}$$

Equations (A.1) and (A.2) are together called the *Fourier transform pair**. These equations can be discretised to represent images, which are themselves digitised representations of continuous objects. The discrete Fourier transform of an image (with dimensions $M \times N$) is given by:

$$F(u,v) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)e^{-i2\pi(ux/M+vy/N)} \tag{A.3}$$

Similarly, the inverse discrete Fourier transform is given by:

$$f(x,y) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} F(u,v)e^{i2\pi(ux/M+vy/N)} \tag{A.4}$$

*The factor $1/2\pi$ in equation A.1 can alternatively placed on the RHS of equation A.2. Another equivalent representation is to include the factor $1/\sqrt{2\pi}$ in both equations.

Equations (A.3) and (A.4) comprise the 2-D discrete Fourier transform (DFT) pair[†] . In all of the above equations u, v represent the frequency domain variables and x, y are the spatial domain variables. Samples in the frequency and spatial domains are related by the following:

$$\Delta u = \frac{1}{M\Delta x} \tag{A.5}$$

and

$$\Delta v = \frac{1}{N\Delta y} \tag{A.6}$$

[†]Corresponding arguments to those given for factor $1/2\pi$ in the continuous DFT apply to the location of the factor $1/MN$ in equations A.3) and (A.4

APPENDIX B

Ill-posed inverse problems

B.1 Conditions leading to a singular problem

It was stated in section 3.1 that image restoration problems can be singular, i.e., exact recovery of the true image distribution from the observed image is not possible. We examine conditions leading to a singular problem here.

We note that two functions are orthogonal over the interval $[a, b]$ if

$$\int_a^b \int_a^b h(\alpha, \beta) x(\alpha, \beta)\, d\alpha d\beta = 0 \tag{B.1}$$

If the function h has parameters i, j then $h(\alpha, \beta, i, j)$ and $x(\alpha, \beta)$ are orthogonal over the interval $[a, b]$ if:

$$\int_a^b \int_a^b h(\alpha, \beta, i, j) x(\alpha, \beta)\, d\alpha d\beta = 0 \quad \forall i, j \tag{B.2}$$

The function h is termed singular with respect to x in this case and restoration is not possible. If the true image distribution comprises of a function x which is orthogonal to h and a function $\tilde{x}$ which is not, then:

$$\int_a^b \int_a^b h(\alpha, \beta, i, j)[x(\alpha, \beta) + \tilde{x}(\alpha, \beta)]\, d\alpha d\beta = \int_a^b \int_a^b h(\alpha, \beta, i, j)\tilde{x}(\alpha, \beta)\, d\alpha d\beta \tag{B.3}$$

The existence of the orthogonal component cannot be inferred from the observed image. Image restoration problems can often be singular.

B.2 Demonstration that image restoration is ill-conditioned

It was asserted in section 3.1 that the problem of image restoration was an ill-posed problem. The ill-conditioned behaviour of the restoration process was demonstrated by Phillips (1962) through use of the Riemann-Lebesgue lemma:

$$\lim_{\phi\to\infty}\lim_{\theta\to\infty}\int_a^b\int_a^b h(\alpha,\beta)\sin(\theta\alpha)\sin(\phi\beta)\,d\alpha d\beta = 0 \tag{B.4}$$

where $h(\theta,\phi)$ is an integrable function. If we have a function h with parameters i,j then equation (B.4) can be written:

$$\lim_{\phi\to\infty}\lim_{\theta\to\infty}\int_a^b\int_a^b h(\alpha,\beta,i,j)\sin(\theta\alpha)\sin(\phi\beta)\,d\alpha d\beta = 0 \tag{B.5}$$

For a linear image receptor with no inherent system noise we have:

$$d(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\alpha,\beta,i,j)x(i,j)\,d\alpha d\beta \tag{B.6}$$

Consider equation (B.5). We can add a sinusoid of infinite frequency to the true image distribution and investigate the effect on the observed image data:

$$\hat{d}(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\alpha,\beta,i,j)[x(i,j)+\sin(\theta\alpha)\sin(\phi\beta)]\,d\alpha d\beta \tag{B.7}$$

Multiplying out we obtain:

$$\hat{d}(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}[h(\alpha,\beta,i,j)x(i,j)+h(\alpha,\beta,i,j)\sin(\theta\alpha)\sin(\phi\beta)]\,d\alpha d\beta \tag{B.8}$$

But by the Riemann-Lebesgue lemma given in equation (B.5) the second term on the RHS is zero and the equation reduces to:

$$\hat{d}(i,j) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\alpha,\beta,i,j)x(i,j)\,d\alpha d\beta = d(i,j) \tag{B.9}$$

So adding an infinite frequency sinusoid to the true image distribution has no effect on the observed image.

APPENDIX C

Maximum Entropy

C.1 Minimization of χ^2

The degree of misfit between observed data d and mock data $\hat{d}$ is quantified by the χ^2 statistic:

$$\chi^2 = \sum_i \frac{(d_i - \hat{d}_i)^2}{\sigma_i^2} \tag{C.1}$$

where σ_i^2 is the noise variance (assumed to be Gaussian) associated with each pixel in the image. The noise (all the values of σ_i^2) will vary across the image and is not known exactly in practice. It is also a combination of sources such as quantum noise and receptor noise. A noise map could perhaps be estimated based on knowledge of the radiation dose, attenuation properties of the object. A more pragmatic approach might be to take as an average value for the whole image, a measurement from within some suitable sub-region of the image. If the noise is taken to be constant across the image then the above equation simplifies to:

$$\chi^2 = \frac{\sum_i (d_i - \hat{d}_i)^2}{\sigma^2} \tag{C.2}$$

The simplifying assumption of constant noise was employed for many of the MEM reconstructions in this work where such an approximation is arguably valid. The mock data is generated from $\hat{d}_i = \sum_j A_{ij}\, x_j$. Let C be the function we wish to minimise:

$$C = \chi^2 = \frac{\sum_i (d_i - \sum_j A_{ij}\, x_j)^2}{\sigma^2} \tag{C.3}$$

To find the minimum, the strategy is therefore to differentiate equation (C.3) and set the derivative to zero. Differentiating, we obtain:

$$\begin{aligned}
\frac{\partial C}{\partial x_k} &= (2/\sigma^2) \sum_i (d_i - \sum_j A_{ij}\, x_j) \left(-A_{ij} \frac{\partial x_j}{\partial x_k} \right) \\
&= (2/\sigma^2) \sum_i (d_i - \sum_j A_{ij}\, x_j) \left(-A_{ij} \delta_{jk} \right) \\
&= (-2/\sigma^2) \sum_i (d_i - \sum_j A_{ij}\, x_j) A_{ik}
\end{aligned} \tag{C.4}$$

So for a minimum we require:

$$\frac{\partial C}{\partial x_k} = 0 \tag{C.5}$$

Therefore

$$\begin{aligned}
(-2/\sigma^2) \sum_i (d_i - \sum_j A_{ij}\, x_j) A_{ik} &= 0 \\
\sum_i (d_i - \sum_j A_{ij}\, x_j) A_{ik} &= 0
\end{aligned} \tag{C.6}$$

The term $d_i - \sum_j A_{ij}\, x_j$ is simply a vector v_i and so we have:

$$\sum_i v_i A_{ik} = 0 \tag{C.7}$$

The matrix A is non-trivial so we have:

$$\begin{aligned}
v_i &= 0 \\
d_i - \sum_j A_{ij}\, x_j &= 0
\end{aligned} \tag{C.8}$$

Minimising χ^2 is therefore equivalent to solving equation (C.8). In matrix notation this is:

$$\begin{aligned}
A\mathbf{x} &= \mathbf{d} \\
A^{-1}A\mathbf{x} &= A^{-1}\mathbf{d} \\
x &= A^{-1}\mathbf{d}
\end{aligned} \tag{C.9}$$

C.2 Deriving the gradient and Hessian of Q

C.2.1 Derivatives of the entropy S

The entropy associated with the reconstructed image $\hat{x}$ is given by:

$$S(\hat{x}) = \sum_i [\hat{x}_i - m_i - \hat{x}_i \ln(\hat{x}_i/m_i)] \tag{C.10}$$

The gradient is obtained by differentiating this expression with respect to x_k:

$$\frac{\partial S}{\partial \hat{x}_i} = \ln m_i - \ln x_i \tag{C.11}$$

The second derivative is obtained by differentiating again, with respect to x_j:

$$\frac{\partial^2 S}{\partial \hat{x}_i \partial \hat{x}_j} = -\frac{\delta_{jk}}{x_j} \tag{C.12}$$

C.2.2 Derivatives of χ^2

Let d_i denote the observed data and $\hat{x}_i$ the estimate of the true (hidden) image distribution. In this analysis we combine the spatially invariant ICF and PSF into a single blurring function $A = c * h$. Mock data $\hat{d}_i$ is generated by:

$$\hat{d}_i = \sum_j A_{j-i} \hat{x}_j \tag{C.13}$$

The degree of misfit between the mock data and observed data is quantified by the χ^2 statistic:

$$\chi^2 = \frac{1}{\sigma^2} \sum_i (d_i - \hat{d}_i)^2 \tag{C.14}$$

where σ^2 is the noise variance, taken to be constant across the image. To obtain the gradient of this expression we will need some other partial derivatives. Differentiating equation C.14 with respect to $\hat{d}_k$ gives:

$$\begin{aligned}\frac{\partial\chi^2}{\partial\hat{d}_k} &= -\frac{1}{\sigma^2}\sum_i 2(d_i-\hat{d}_i)\frac{\partial\hat{d}_i}{\partial\hat{d}_k} \\ &= -\frac{1}{\sigma^2}\sum_i 2(d_i-\hat{d}_i)\delta_{ik} \\ &= -\frac{2}{\sigma^2}\sum_k (d_k-\hat{d}_k) \end{aligned} \tag{C.15}$$

Differentiating equation C.13 with respect to $\hat{x}_k$ gives:

$$\begin{aligned}\frac{\partial\hat{d}_i}{\partial\hat{x}_k} &= \sum_j A_{j-i}\delta_{jk} \\ &= A_{k-i} \end{aligned} \tag{C.16}$$

We now have the information required to form the gradient of equation C.14:

$$\begin{aligned}\frac{\partial\chi^2}{\partial\hat{x}_k} &= \sum_i \frac{\partial\chi^2}{\partial\hat{d}_i}\frac{\partial\hat{d}_i}{\partial\hat{x}_k} \\ &= -\frac{2}{\sigma^2}\sum_i (d_i-\hat{d}_i)A_{k-i} \\ &= -\frac{2}{\sigma^2}\sum_i A^T_{i-k}(d_i-\hat{d}_i) \end{aligned} \tag{C.17}$$

The gradient of χ^2 is therefore the convolution of A^T with $\mathbf{d}-\hat{\mathbf{d}}$. The second derivative is obtained by differentiating equation C.17 with respect to $\hat{x}_j$.

$$\frac{\partial^2\chi^2}{\partial\hat{x}_j\partial\hat{x}_k} = \frac{2}{\sigma^2}\sum_i A_{j-i}A_{k-i} \tag{C.18}$$

APPENDIX D

Matlab codes

D.1 High pass and lowpass filters

```
function [f,G,H]=imhlpass(I,D0,HORL,filtertype,varargin)
% [f,G,H]=imhlpass(I,D0,'HP or LP','filtertype',bn)
%
% Highpass and lowpass frequency filtering.
%
% The three types of filter are:
%
% (1) Ideal
% (2) Butterworth
% (3) Gaussian
%
% The input arguments depend on the type of filter being employed.  There are a few parameters
% common to all filters:
%
% D0 is the cutoff frequency
% 'LP' and 'HP' represents lowpass and highpass respectively
%
% The output arguments are the same for each filter.
%
% f is the filtered image
% G is the transfer function of the filtered image
% H is the transfer function of the filter
%
% Ideal filter
% [f,G,H]=imhlpass(I,D0,'LP','ideal')              Ideal lowpass
% [f,G,H]=imhlpass(I,D0,'HP','ideal')              Ideal highpass
%
% Butterworth filter
% [f,G,H]=imhlpass(I,D0,'LP','butterworth',bn)     Butterworth lowpass
% [f,G,H]=imhlpass(I,D0,'HP','butterworth',bn)     Butterworth highpass
%
% bn represents the order of the filter.
%
```

```
% Gaussian filter
% [f,G,H]=imhlpass(I,D0,'LP','gaussian')            Gaussian lowpass
% [f,G,H]=imhlpass(I,D0,'HP','gaussian')            Gaussian highpass
%
% The input image can be of type UINT8, UINT16 or DOUBLE.  The output image has the same class
% as the input image.
%
% These are implementations of the filters in "Digital image processing (2nd Ed)"
% by Gonzalez and Wintz.

% Check the number of arguments
error(nargchk(4,5,nargin))

% Image dimensions
[m,n]=size(I);

% DFT of the image
F=fft2(I);

% Form a mesh for the transfer function
[v,u]=meshgrid(1:n,1:m);

if nargin==5
    % Butterworth filter
    bn=varargin{1};
    switch HORL
    case {'LP'}

D=sqrt((u-m/2).^2+(v-n/2).^2);
H=1./(1+(D./D0).^(2*bn));

    case {'HP'}
F=fft2(I);

% Design the filter
[v,u]=meshgrid(1:n,1:m);

D=sqrt((u-m/2).^2+(v-n/2).^2);
H=1./(1+(D0./D).^(2*bn));
    end

else

    switch filtertype
    case {'ideal'}
        switch HORL
        case {'LP'}
            D=sqrt((u-m/2).^2+(v-n/2).^2);
H=ones(m,n);
HI=find(D<D0); HO=find(D>D0);
H(HI)=1;
H(HO)=0;
        case {'HP'}
D=sqrt((u-m/2).^2+(v-n/2).^2);
H=ones(m,n);
HI=find(D<D0); HO=find(D>D0);
H(HI)=0;
H(HO)=1;
        end

    case {'gaussian'}
        switch HORL
```

```
        case {'LP'}
            D=sqrt((u-m/2).^2+(v-n/2).^2);
            H=exp((D.^2)./(-2*(D0^2)));
        case {'HP'}
            D=sqrt((u-m/2).^2+(v-n/2).^2);
            H=1-exp((D.^2)./(-2*(D0^2)));
        end
    end

end

% Shift the transform and multiply with original image transform
H=fftshift(H);
G=F.*H;

% Convert filtered image to spatial domain and output
f=real(ifft2(G));

% Make output type same as input type
% 8-bit
if isa(I,'uint8')==1
    f=uint8(round(f-1));
end
% 16-bit
if isa(I,'uint16')==1
    f=uint16(round(f-1));
end
% % double
% if isa(I,'double')==1
%     f=im2double(f);
% end
```

D.2 High frequency emphasis filter

```
function [g,H]=imhfe(f,a,b,D0,filtertype,varargin)
% [g,H]=imhfe(f,a,b,D0,'filtertype',bn)
%
% will form a high frequency emphasis (HFE) filter transfer function
% and apply it to an image.  The HFE filter modifies the transfer function
% of a highpass filter.  The three types of filter are:
%
% (1) Ideal
% (2) Butterworth (must supply additional argument bn (1,2,3,...etc)
% (3) Gaussian
%
% f is the input image (type UINT8)
% HP is the transfer function of the highpass filter
% a,b are the parameters which define the HFE function according to:
%
% H(u,v) = a + b*HP(u,v)
%
% where a>=0  (typically, 0.25 < a < 0.5)
% and   b>a   (typically, 1.5 < b < 2.0)
%
% g is the output image (type UINT8)
% H is the transfer function of the HFE
%
% Use IMADJUST or HISTEQ to adjust the histogram of the output image if necessary

% Form the specified high pass filter HP
if nargin==6
    bn=varargin{1};
    [hp,G,HPass]=imhlpass(f,D0,'HP',filtertype,bn);
else
    [hp,G,HPass]=imhlpass(f,D0,'HP',filtertype);
end

% Form the high frequency emphasis filter
H=a+b*HPass;

F=fft2(f);

Gout=H.*F;

% Convert filtered image to spatial domain and output
g=real(ifft2(Gout));

% Make output type same as input type
% 8-bit
if isa(f,'uint8')==1
    g=uint8(round(g-1));
end
% 16-bit
if isa(f,'uint16')==1
    g=uint16(round(g-1));
end
```

D.3 Inverse and pseudoinverse filters

```
function [f,H]=impinverse(I,PSF,lambda)
% [f,H]=impinverse(I,PSF,lambda)
%
% This function will restore an image using an inverse filter.
%
% I is the degraded input image (type UINT8).
% h is the PSF matrix.
% lambda is the threshold value, chosen to avoid inverting zeros in the frequency domain.
%
% f is the restored image (type UINT8)
% H is the transfer function of the inverse filter
%
% Note:  With lambda = 0, the filter is a straightforward inverse filter
%        With lambda > 0, the filter becomes a pseudoinverse filter
%
%        Choose lambda > 0 for noisy images.
%
% See also IMWIENER

[m,n]=size(I);    % Image dimensions
[mm,nn]=size(PSF);  % PSF dimensions

pm=m+mm; pn=n+nn; % Calculate padsize

% Calculate the FFT of the padded image and padded PSF
H=psf2otf(PSF,[pm pn]);
G=fft2(I,pm,pn);

% Replace values in 1/H which are < lambda in abs(H)
Hlambda=find(abs(H)<lambda);
HI=1./H;
HI(Hlambda)=lambda;

% Calculate the transform F of the inverse filtered image
F=HI.*G;

% Convert back to spatial domain
f=ifft2(F,pm,pn);
f=real(f); f=uint8(round(f-1));

% Crop the output image to the same dimensions as the input image
window=[0 0 n m];
f=imcrop(f,window);
```

D.4 Wiener filter

```
function [f,HW]=imwiener(g,h,K)
% [f,HW]=imwiener(g,h,K)
%
% This function will restore an image using a Wiener filter.  The input can be
% a grayscale or RGB colour image.
%
% g is the degraded input image (type DOUBLE)
% h is the PSF matrix
% K is the constant term - an approximation to 1/signal-to-noise ratio
%
% f is the restored image
% HW is the Wiener filter transfer function
%
% Setting K=0 reduces this to an inverse filter.

% Blur the image edges to reduce ringing effects
g=edgetaper(g,h);

% Store the image type of the degraded image
gtemp=g(1);

[m,n,ch]=size(g);    % Image dimensions
[mm,nn]=size(h);  % PSF dimensions

pm=m+mm; pn=n+nn; % Calculate padsize

g=im2double(g);

% Is this a colour image (3 channels)?
flag=isrgb(g);

% Calculate the FFT  padded PSF
H=psf2otf(h,[pm pn]);

% Calculate the transform F of the Wiener filtered image
HSTAR=conj(H); HP=HSTAR.*H;
HW=(HSTAR./(HP+K));

for its=1:ch  % Calculate the FFT of each channel of the padded image
    G(:,:,its)=fft2(g(:,:,its),pm,pn);
F(:,:,its)=HW.*G(:,:,its);
end

% Convert back to spatial domain
f=ifft2(F); f=real(f);

% Convert restored image to same type as original (d)
if isa(gtemp,'uint8')==1
    f=im2uint8(f);
end
if isa(gtemp,'uint16')==1
    f=im2uint16(f);
end
if isa(gtemp,'double')==1
    f=im2double(f);
end

% Crop the output image to the same dimensions as the input image
window=[0 0 n m]; f=imcrop(f,window);
```

D.5 Maximum entropy

```
function [v,x,CR,alpha,F,res]=immem(d,A,sigsq,b,ICFW)
% IMMEM.m will restore an image using Maximum Entropy Method.
%
% [v,x,CR,alpha,F,res]=IMMEM(d,A,sigsq,b,ICFW);
%
% d === Data/degraded image (type UINT8 or UINT16)
% A === Normalised PSF
% sigsq === noise variance (sigma^2)
% b === the default (flat) image value
% ICFW === Width of the intrinsic correlation function (Use ICFW=1 for no pixel correlations).
%
% v === the visible (maxent) image (same type as d)
% x === the estimate of the hidden image (same type as d)
% CR === ratio of Chi_Restored to Chi_Expected
% alpha === final value of the parameter giving weight to the entropy
% F === Movie frames for intermediate stages of the restoration

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

tic; % Start timer

dtemp=d(1); % Store the image type of the degraded image

[m,n]=size(d); dim=size(d);   % Image dimensions (m,n). Stored in dim.
[mm,nn]=size(A); N=prod(dim); NN=mm*nn;  % PSF dimensions

TolChi=0.1; % Termination criterion for CR

ICF=fspecial('gaussian',ICFW,1); % Gaussian ICF (user-specified size)

PSF2=conv2(A,ICF,'full');

% Blur the image edges to reduce edge ringing effects in the restored image
d=edgetaper(d,A);

% Convert images to DOUBLE and scale to range [0 1] and convert to vectors
d=im2double(d);
d=d(:);

% Optimisation parameters
lb=(1/1024)*ones(N,1); ub=ones(N,1); % Upper and Lower bounds on the solution
alpha=1e+3; % Initial value of alpha
x0=b*ones(N,1); % Initial guess at a solution

options=optimset('MaxPCGIter',5,'MaxFunEvals',50,'GradObj','on','Hessian','on','Display','final');
disp(' ');
s1 = sprintf('   its   alpha      Chi^2/N');
disp(s1)
CR=inf; its=1;

while abs(CR-1)>TolChi

% Optimisation (x is the vector of restored image data)
    x=fmincon(@entropy_fun,x0,[],[],[],[],lb,[],[],options,A,PSF2,d,b,m,n,N,sigsq,alpha,dim);
    x0=x;

% Calculate Chi-squared Statistic
    y=mockdata(x,PSF2,dim);
```

```
  CR=chi2(d,y,sigsq); CR=CR/N; % Chi^2 relative to expected value of N

    % Increase or decrease alpha for next round of iterations
    oldalpha=alpha;
    if CR>1
        dalpha=-alpha*((CR-1)/CR);
    else
        dalpha=alpha*(1-CR);
    end
    if abs(CR-1)<2.5
        options=optimset('MaxFunEvals',50,'GradObj','on','Hessian','on','Display','final');
    end

    alpha=alpha+dalpha; % New value of alpha

    % Generate movie frames and store in F
    x1=reshape(x,dim);
    figure,imshow(x1,'notruesize')
    F(its)=getframe;
    close

    its=its+1;
    pack;
    s2 = sprintf('\t %d \t %0.5g \t %0.5g',its-1,oldalpha,CR);
    disp(s2)

end

x=reshape(x,dim);
v=conv2(x,ICF,'same');

% Final movie frame with restored image
figure,imshow(v,'notruesize')
F(its)=getframe;
close

% Residuals
res=imabsdiff(reshape(d,dim),v);

time_elapsed=toc

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [Q,dQ,d2Q]=entropy_fun(x,A,PSF2,d,b,m,n,N,sigsq,alpha,dim)

% Entropy (Skilling 1988) and derivatives
S = sum(x - b - (x.*(log(x/b))));
DS=1 - log(x/b);
D2S=spdiags((1./(1./x)),0,N,N);

% Trial restoration
y=mockdata(x,PSF2,dim);

% Chi^2 and derivatives
C=chi2(d,y,sigsq);
DTC=reshape(d-y,dim); DTC2=conv2(DTC,A','same');
DC=(-2/sigsq)*DTC2(:);

% Objective function, gradient and approximate Hessian

Q   = -alpha*S+C;
dQ  = -alpha*DS+DC;
```

```
d2Q = -alpha*D2S;

%%%
```

D.6 Maximum entropy: 3-plane tomographic reconstruction

```
function [x1,x2,x3,CR,alpha]=me_tomo3(d1,d2,d3,sigsq)
% [v1,v2,v3,x1,x2,x3,CR,alpha]=me_tomo3(d1,d2,d3,A1,A2,A3,sigsq)
%
% me_tomo3.m will reconstruct 3 planes using MEM.
%
% di is the degraded image
% Ai is the normalised PSF
% sigsq is the noise variance (sigma^2)
% b is the default image (same type as d)
%
% vi is the estimate of the visible (maxent) image (same type as d)
% xi is the estimate of the hidden image (same type as d)
% CR is the ratio of Chi_Restored to Chi_Expected
% alpha is the final value of the parameter giving weight to the entropy
%
% The entropy is maximised by using fmincon to minimise the function:
% alpha*S + C where
%
% C = (1/sigsq)*sum((d-A*x).^2) and S = sum(x.*(log(x./b)-1))

% Start timer
tic;

global A1 A2 A3 B1 B2 B3 C1 C2 C3

% A1=1; A2=ones(1,91)/91; A3=ones(1,193)/193;
% B1=ones(1,173)/173; B2=1; B3=ones(1,98)/98;
% C1=ones(1,261)/261; C2=ones(1,93)/93; C3=1;

% Sim stacks
A1=1; A2=ones(1,10)/10; A3=ones(1,20)/20;
B1=ones(1,10)/10; B2=1; B3=ones(1,10)/10;
C1=ones(1,20)/20; C2=ones(1,10)/10; C3=1;

% Store the image type of the degraded image
dtemp=d1(1);

[m,n]=size(d1);   % Image dimensions
dim=size(d1);
N=m*n;

% Default image: Uniform, low intensity background
b=0.005;

TolChi=0.25;

d1=d1(:); d2=d2(:); d3=d3(:);

% Optimisation parameters
options=optimset('TolX',0.005,'MaxFunEvals',30,'GradObj','on','Hessian','on','Display','iter');
disp(' ');
s1 = sprintf('   its    alpha      Chi^2/N');
disp(s1)

lb=zeros(3*N,1); ub=ones(3*N,1); % Upper and Lower bounds on the solution
alpha=250; % Initial value of alpha
```

```
x0=b*ones(3*N,1); % Initial guess at a solution

its=0;
CR=inf;

while its<5
%while abs(CR-1)>TolChi

% Optimisation (x is the vector of restored image data)
    x=fmincon(@mem_tomo3_fun,x0,[],[],[],[],lb,ub,[],options,d1,d2,d3,m,n,N,sigsq,alpha,b,dim);
    x0=x;

    x1=x([1:N]); x1=reshape(x1,m,n);
    x2=x([N+1:2*N]); x2=reshape(x2,m,n);
    x3=x([2*N+1:3*N]); x3=reshape(x3,m,n);

    % Calculate global Chi^2 statistic
    y1=tomo_trialres(x,A1,A2,A3,dim,N);
    [CR1,C1temp]=tomo3_chi2(d1,y1,sigsq);

    y2=tomo_trialres(x,B1,B2,B3,dim,N);
    [CR2,C2temp]=tomo3_chi2(d2,y2,sigsq);

    y3=tomo_trialres(x,C1,C2,C3,dim,N);
    [CR3,C3temp]=tomo3_chi2(d3,y3,sigsq);

    Ctemp=[C1temp;C2temp;C3temp];
    C=(1/sigsq)*sum(Ctemp.^2);

    CR=C/N;

    % Increase or decrease alpha for next round of iterations
    oldalpha=alpha;
    if CR>1
        dalpha=-1*(alpha*(CR-1)/CR);
    else
        dalpha=alpha*(1-CR);
    end
    if abs(CR-1)<2.5
        options=optimset('TolX',0.001,'MaxFunEvals',30,'GradObj','on','Hessian','on','Display','iter');
    end
    alpha=alpha+dalpha;

    its=its+1;
    pack;
    s2 = sprintf('\t %d \t %0.5g \t %0.5g',its,oldalpha,CR);
    disp(s2)

end

x1=reshape(x1,dim); x2=reshape(x2,dim); x3=reshape(x3,dim);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [f,g,h]=mem_tomo3_fun(x,d1,d2,d3,m,n,N,sigsq,alpha,b,dim)

global A1 A2 A3 B1 B2 B3 C1 C2 C3

% % Entropy, gradient and Hessian
S = sum(x - b - (x.*(log(x./b))));
DS=1 - log(x./b);
D2S=spdiags((1./(1./x)),0,3*N,3*N);
```

```
% Trial restoration
y1=tomo_trialres(x,A1,A2,A3,dim,N);
y2=tomo_trialres(x,B1,B2,B3,dim,N);
y3=tomo_trialres(x,C1,C2,C3,dim,N);

% Chi-square, gradient and Hessian
C1temp=d1-y1;
C2temp=d2-y2;
C3temp=d3-y3;

Ctemp=[C1temp;C2temp;C3temp];
C=(1/sigsq)*sum(Ctemp.^2);

DC1=reshape(C1temp,m,n);
DC2=reshape(C2temp,m,n);
DC3=reshape(C3temp,m,n);
DC=(-2/sigsq)*[DC1(:); DC2(:); DC3(:)];

% Final functions, gradient and hessian approximations

f=-alpha*S+C;

g=-alpha*DS+DC;

h=-alpha*D2S;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function y=tomo_trialres(x,A1,A2,A3,dim,N)

% Trial restoration.  Generation of mock data from an estimate of the hidden image.

% xi === estimates of the true image data (matrix)
% Ai === PSFs
% y === mock data
% dim === dimensions of the image

x1=reshape(x([1:N]),dim); x2=reshape(x([N+1:2*N]),dim); x3=reshape(x([2*N+1:3*N]),dim);

y1=imfilter(x1,A1,'conv','same','replicate');
y2=imfilter(x2,A2,'conv','same','replicate');
y3=imfilter(x3,A3,'conv','same','replicate');

y=y1+y2+y3;
y=y(:);

%%%%%%%%%%%%%%%%%%%
```

D.7 Maximum entropy: 7-plane tomographic reconstruction

```
function [x1,x2,x3,x4,x5,x6,x7,CR,alpha]=me_tomo7(d1,d2,d3,d4,d5,d6,d7,psfdims)
% [x1,x2,x3,x4,x5,x6,x7,CR,alpha]=me_tomo7(d1,d2,d3,d4,d5,d6,d7,psfdims)
%
% me_tomo7.m will reconstruct 7 planes using MEM.
%
% di is the degraded image
% Ai is the normalised PSF
% sigsq is the noise variance (sigma^2)
% b is the default image (same type as d)
%
% vi is the estimate of the visible (maxent) image (same type as d)
% xi is the estimate of the hidden image (same type as d)
% CR is the ratio of Chi_Restored to Chi_Expected
% alpha is the final value of the parameter giving weight to the entropy
%
% The entropy is maximised by using fmincon to minimise the function:
% alpha*S + C where
%
% C = (1/sigsq)*sum((d-A*x).^2) and S = sum(x.*(log(x./b)-1))

% Start timer
tic;

global A1 A2 A3 A4 A5 A6 A7
global B1 B2 B3 B4 B5 B6 B7
global C1 C2 C3 C4 C5 C6 C7
global D1 D2 D3 D4 D5 D6 D7
global E1 E2 E3 E4 E5 E6 E7
global F1 F2 F3 F4 F5 F6 F7
global G1 G2 G3 G4 G5 G6 G7

d1=(double(d1)+1)./1024;
d2=(double(d2)+1)./1024;
d3=(double(d3)+1)./1024;
d4=(double(d4)+1)./1024;
d5=(double(d5)+1)./1024;
d6=(double(d6)+1)./1024;
d7=(double(d7)+1)./1024;

A1=psfdims(1,1); B2=psfdims(2,2); C3=psfdims(3,3); D4=psfdims(4,4);
E5=psfdims(5,5); F6=psfdims(6,6); G7=psfdims(7,7);

A2=ones(psfdims(1,2),1)/psfdims(1,2); A3=ones(psfdims(1,3),1)/psfdims(1,3);
A4=ones(psfdims(1,4),1)/psfdims(1,4);
A5=ones(psfdims(1,5),1)/psfdims(1,5); A6=ones(psfdims(1,6),1)/psfdims(1,6);
A7=ones(psfdims(1,7),1)/psfdims(1,7);

B1=ones(psfdims(2,1),1)/psfdims(2,1); B3=ones(psfdims(2,3),1)/psfdims(2,3);
B4=ones(psfdims(2,4),1)/psfdims(2,4);
B5=ones(psfdims(2,5),1)/psfdims(2,5); B6=ones(psfdims(2,6),1)/psfdims(2,6);
B7=ones(psfdims(2,7),1)/psfdims(2,7);

C1=ones(psfdims(3,1),1)/psfdims(3,1); C2=ones(psfdims(3,2),1)/psfdims(3,2);
C4=ones(psfdims(3,4),1)/psfdims(3,4);
C5=ones(psfdims(3,5),1)/psfdims(3,5); C6=ones(psfdims(3,6),1)/psfdims(3,6);
C7=ones(psfdims(3,7),1)/psfdims(3,7);
```

```
D1=ones(psfdims(4,1),1)/psfdims(4,1); D2=ones(psfdims(4,2),1)/psfdims(4,2);
D3=ones(psfdims(4,3),1)/psfdims(4,3);
D5=ones(psfdims(4,5),1)/psfdims(4,5); D6=ones(psfdims(4,6),1)/psfdims(4,6);
D7=ones(psfdims(4,7),1)/psfdims(4,7);

E1=ones(psfdims(5,1),1)/psfdims(5,1); E2=ones(psfdims(5,2),1)/psfdims(5,2);
E3=ones(psfdims(5,3),1)/psfdims(5,3);
E4=ones(psfdims(5,4),1)/psfdims(5,4); E6=ones(psfdims(5,6),1)/psfdims(5,6);
E7=ones(psfdims(5,7),1)/psfdims(5,7);

F1=ones(psfdims(6,1),1)/psfdims(6,1); F2=ones(psfdims(6,2),1)/psfdims(6,2);
F3=ones(psfdims(6,3),1)/psfdims(6,3);
F4=ones(psfdims(6,4),1)/psfdims(6,4); F5=ones(psfdims(6,5),1)/psfdims(6,5);
F7=ones(psfdims(6,7),1)/psfdims(6,7);

G1=ones(psfdims(7,1),1)/psfdims(7,1); G2=ones(psfdims(7,2),1)/psfdims(7,2);
G3=ones(psfdims(7,3),1)/psfdims(7,3);
G4=ones(psfdims(7,4),1)/psfdims(7,4); G5=ones(psfdims(7,5),1)/psfdims(7,5);
G6=ones(psfdims(7,6),1)/psfdims(7,6);

[m,n]=size(d1);   % Image dimensions
dim=size(d1);
N=m*n;

% Default image: Uniform, low intensity background.
b=(1/2000);  TolChi=0.5;

d1=mat2gray(10.^(d1.*1.6)); d1=d1(:); s1=0.001*sqrt(1-d1)+0.0001;
d2=mat2gray(10.^(d2.*1.6)); d2=d2(:); s2=0.001*sqrt(1-d2)+0.0001;
d3=mat2gray(10.^(d3.*1.6)); d3=d3(:); s3=0.001*sqrt(1-d3)+0.0001;
d4=mat2gray(10.^(d4.*1.6)); d4=d4(:); s4=0.001*sqrt(1-d4)+0.0001;
d5=mat2gray(10.^(d5.*1.6)); d5=d5(:); s5=0.001*sqrt(1-d5)+0.0001;
d6=mat2gray(10.^(d6.*1.6)); d6=d6(:); s6=0.001*sqrt(1-d6)+0.0001;
d7=mat2gray(10.^(d7.*1.6)); d7=d7(:); s7=0.001*sqrt(1-d7)+0.0001;

% Noise maps
sigsq=[s1 s2 s3 s4 s5 s6 s7];

% Optimisation parameters
options=optimset('TolFun',1e-4,'MaxFunEvals',35,'GradObj','on','Hessian','on','Display','iter');
disp(' ');
t1 = sprintf('   its   alpha      Chi^2/N');
disp(t1)

lb=0.001*ones(7*N,1); ub=ones(7*N,1); % Upper and Lower bounds on the solution
alpha=0.001; % Initial value of alpha
x0=b*ones(7*N,1); % Initial guess at a solution

its=1;
CR=inf;
while its<2
% while abs(CR-1)>TolChi

% Optimisation (x is the vector of restored image data)
    x=fmincon(@mem_tomo_fun7,x0,[],[],[],[],lb,ub,[],options,d1,d2,d3,d4,d5,d6,d7,m,n,N,sigsq,alpha,b,dim);
    x1=x([1:N]); x1=reshape(x1,m,n);
    x2=x([N+1:2*N]); x2=reshape(x2,m,n);
    x3=x([2*N+1:3*N]); x3=reshape(x3,m,n);
    x4=x([3*N+1:4*N]); x4=reshape(x4,m,n);
    x5=x([4*N+1:5*N]); x5=reshape(x5,m,n);
    x6=x([5*N+1:6*N]); x6=reshape(x6,m,n);
```

```
    x7=x([6*N+1:7*N]); x7=reshape(x7,m,n);

    % Calculate global Chi^2 statistic
    y1=tomo_trialres7(x,A1,A2,A3,A4,A5,A6,A7,dim,N,m,n);
    [CR1,C1temp]=tomo7_chi2(d1,y1,s1);

    y2=tomo_trialres7(x,B1,B2,B3,B4,B5,B6,B7,dim,N,m,n);
    [CR2,C2temp]=tomo7_chi2(d2,y2,s2);

    y3=tomo_trialres7(x,C1,C2,C3,C4,C5,C6,C7,dim,N,m,n);
    [CR3,C3temp]=tomo7_chi2(d3,y3,s3);

    y4=tomo_trialres7(x,D1,D2,D3,D4,D5,D6,D7,dim,N,m,n);
    [CR4,C4temp]=tomo7_chi2(d4,y4,s4);

    y5=tomo_trialres7(x,E1,E2,E3,E4,E5,E6,E7,dim,N,m,n);
    [CR5,C5temp]=tomo7_chi2(d5,y5,s5);

    y6=tomo_trialres7(x,F1,F2,F3,F4,F5,F6,F7,dim,N,m,n);
    [CR6,C6temp]=tomo7_chi2(d6,y6,s6);

    y7=tomo_trialres7(x,G1,G2,G3,G4,G5,G6,G7,dim,N,m,n);
    [CR7,C7temp]=tomo7_chi2(d7,y7,s7);

    C=CR1+CR2+CR3+CR4+CR5+CR6+CR7;
    CR=C/(7*N);

    % Increase or decrease alpha for next round of iterations
    oldalpha=alpha;
    if CR>1
        dalpha=-1*(alpha*(CR-1)/CR);
    else
        dalpha=alpha*(1-CR);
    end
    if abs(CR-1)<2.5
        options=optimset('MaxFunEvals',35,'GradObj','on','Hessian','on','Display','iter');
    end
    alpha=alpha+dalpha;

    its=its+1;
    pack;
    t2 = sprintf('\t %d \t %0.5g \t %0.5g',its,oldalpha,CR);
    disp(t2)

end

x1=reshape(x1,dim); x2=reshape(x2,dim); x3=reshape(x3,dim);
x4=reshape(x4,dim); x5=reshape(x5,dim); x6=reshape(x6,dim); x7=reshape(x7,dim);

toc

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [f,g,h]=mem_tomo_fun7(x,d1,d2,d3,d4,d5,d6,d7,m,n,N,sigsq,alpha,b,dim)

global A1 A2 A3 A4 A5 A6 A7
global B1 B2 B3 B4 B5 B6 B7
global C1 C2 C3 C4 C5 C6 C7
global D1 D2 D3 D4 D5 D6 D7
global E1 E2 E3 E4 E5 E6 E7
global F1 F2 F3 F4 F5 F6 F7
global G1 G2 G3 G4 G5 G6 G7
```

```
% Entropy, gradient and Hessian
S = sum(x - b - (x.*(log(x./b))));
DS=1 - log(x./b);
D2S=spdiags((1./(1./x)),0,7*N,7*N);

% Trial restoration
y1=tomo_trialres7(x,A1,A2,A3,A4,A5,A6,A7,dim,N,m,n);
y2=tomo_trialres7(x,B1,B2,B3,B4,B5,B6,B7,dim,N,m,n);
y3=tomo_trialres7(x,C1,C2,C3,C4,C5,C6,C7,dim,N,m,n);
y4=tomo_trialres7(x,D1,D2,D3,D4,D5,D6,D7,dim,N,m,n);
y5=tomo_trialres7(x,E1,E2,E3,E4,E5,E6,E7,dim,N,m,n);
y6=tomo_trialres7(x,F1,F2,F3,F4,F5,F6,F7,dim,N,m,n);
y7=tomo_trialres7(x,G1,G2,G3,G4,G5,G6,G7,dim,N,m,n);

% Chi-square, gradient and Hessian
C1temp=sum(((d1-y1).^2)./sigsq(:,1));
C2temp=sum(((d2-y2).^2)./sigsq(:,2));
C3temp=sum(((d3-y3).^2)./sigsq(:,3));
C4temp=sum(((d4-y4).^2)./sigsq(:,4));
C5temp=sum(((d5-y5).^2)./sigsq(:,5));
C6temp=sum(((d6-y6).^2)./sigsq(:,6));
C7temp=sum(((d7-y7).^2)./sigsq(:,7));

Ctemp=[C1temp;C2temp;C3temp;C4temp;C5temp;C6temp;C7temp];
C=sum(Ctemp);

DC1=reshape(d1-y1,m,n); DTC1=-2*imfilter(DC1,A1','conv','same',0);
DC2=reshape(d2-y2,m,n); DTC2=-2*imfilter(DC2,B2','conv','same',0);
DC3=reshape(d3-y3,m,n); DTC3=-2*imfilter(DC3,C3','conv','same',0);
DC4=reshape(d4-y4,m,n); DTC4=-2*imfilter(DC4,D4','conv','same',0);
DC5=reshape(d5-y5,m,n); DTC5=-2*imfilter(DC5,E5','conv','same',0);
DC6=reshape(d6-y6,m,n); DTC6=-2*imfilter(DC6,F6','conv','same',0);
DC7=reshape(d7-y7,m,n); DTC7=-2*imfilter(DC7,G7','conv','same',0);

DC=[DTC1(:); DTC2(:); DTC3(:); DTC4(:); DTC5(:); DTC6(:); DTC7(:)];

% Final functions, gradient and hessian approximations

f=-alpha*S+C;

g=-alpha*DS+DC;

h=-alpha*D2S;

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function [C,Ctemp]=tomo7_chi2(d,tr,sigma2)

% Calculate the Chi-Squared statistic for observed and trial data for 3 tomo images
%
% d === the observed image data (vector)
% tr  === the trial/mock data (vector)
% s2  === the noise variance vector (sigma^2)

Ctemp=d-tr;
C=sum((Ctemp.^2)./sigma2);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function y=tomo_trialres7(x,A1,A2,A3,A4,A5,A6,A7,dim,N,m,n)
```

```
% Trial restoration.  Generation of mock data from an estimate of the hidden image.

% xi === estimates of the true image data (matrix)
% Ai === PSFs
% y === mock data
% dim === dimensions of the image

x1=reshape(x([1:N]),m,n); x2=reshape(x([N+1:2*N]),m,n);
x3=reshape(x([2*N+1:3*N]),m,n); x4=reshape(x([3*N+1:4*N]),m,n);
x5=reshape(x([4*N+1:5*N]),m,n); x6=reshape(x([5*N+1:6*N]),m,n);
x7=reshape(x([6*N+1:7*N]),m,n);

x1=reshape(x1,dim); x2=reshape(x2,dim); x3=reshape(x3,dim);
x4=reshape(x4,dim); x5=reshape(x5,dim); x6=reshape(x6,dim);
x7=reshape(x7,dim);

y1=imfilter(x1,A1,'conv','same',0); y2=imfilter(x2,A2,'conv','same',0);
y3=imfilter(x3,A3,'conv','same',0); y4=imfilter(x4,A4,'conv','same',0);
y5=imfilter(x5,A5,'conv','same',0); y6=imfilter(x6,A6,'conv','same',0);
y7=imfilter(x7,A7,'conv','same',0);

y=y1+y2+y3+y4+y5+y6+y7;
y=y(:);
```

List of references

Andrews, H. C. and Hunt, B. R. (1977). *Digital image restoration.* Signal processing series. Prentice-Hall.

Badea, C., Kolitsi, Z., and Pallikarakis, N. (1998). A wavelet-based method for removal of out-of-plane structures in digital tomosynthesis. *Computerized Medical Imaging and Graphics*, 22:309–15.

Banham, M. R. and Katsaggelos, A. K. (1997). Digital image restoration. *IEEE Signal Processing Magazine*, pages 24–41.

Barrett, H. H. and Swindell, W. (1981). *Radiological imaging: The theory of image formation, detection and, processing*, volume 1. Harcourt Brace Jovanovich, New York.

Bates, R. H. T. and McDonnell, M. J. (1989). *Image restoration and reconstruction.* Oxford Science Publications.

Baxes, G. A. (1984). *Digital image processing: a practical primer.* Prentice-Hall, Englewood Cliffs, N.J.

Bi, H. and Boerner, G. (1994). When does the Richardson-Lucy deconvolution converge? *Astron. Astrophys. Suppl.*, 108:409–15.

Boardman, A. K. (1979). Constrained optimisation and its application to scintigraphy. *Phys. Med. Biol.*, 24(2):363–71.

Bochud, F. O., Valley, J. F., Verdun, F. R., Hessler, C., and Schnyder, P. (1999). Estimation of the noisy component of anatomical backgrounds. *Med. Phys.*, 26(7):1365–70.

Bonnet, S., Peyrin, F., Turjman, F., and Prost, R. (2002). Multiresolution reconstruction in fan-beam tomography. *IEEE Trans. on Image Processing*, 11(3):169–76.

Bushberg, J. T., Seibert, J. A., Leidholdt, E. M., and Boone, J. M. (2002). *The essential physics of medical imaging.* Lippincott Williams and Wilkins, 2nd edition.

Butterfield, K. J., Dagenais, M., and Clokie, C. (1997). Linear tomography's clinical accuracy and validity for presurgical dental implant analysis. *Oral Surg. Oral Med. Oral Pathol. Oral Radiol. Endod.*, 84(2):203–9.

Calicchia, A., Gambaccini, M., Indovina, P. L., Mazzei, F., and Pugliani, L. (1996). Niobium/molybdenum k-edge filtration in mammography: contrast and dose evaluation. *Phys. Med. Biol.*, 41(9):1717–26.

Carmichael, F. A., Hirschmann, P. N., Scaife, B., Sheard, L., and Mackenzie, A. (2000). A comparison of the diagnostic utility of two image receptors for panoramic radiography. *Dentomaxillofac Radiol*, 29(1):57–60.

Carton, A.-K. (2004). *Development and application of methods for the assessment of image quality and detector performance in digital mammography.* PhD thesis, Katholieke Universiteit Leuven, Faculty of Medicine, Department of Radiology.

Castleman, K. R. (1979). *Digital image processing.* Prentice-Hall signal processing series. Prentice-Hall, Englewood Cliffs, N.J.

Cesar, L. J., Schueler, B. A., Zink, F. E., Daly, T. R., Taubel, J. P., and Jorgenson, L. L. (2001). Artefacts found in computed radiography. *British Journal of Radiology*, 74:195–202.

Chakraborty, D. P. (1999). The effect of the antiscatter grid on full-field digital mammography phantom images. *J. Dig. Imaging*, 12:12–22.

Chakraborty, D. P., Yester, M. V., Barnes, G. T., and Lakshminarayanan, A. V. (1984). Selfmasking subtraction tomosynthesis. *Radiology*, 150:225–229.

Chen, Z. and Ning, R. (2003). Why should breast tumour detection go three dimensional? *Phys. Med. Biol.*, 48:2217–28.

Chevalier, M., Moran, P., Ten, J. I., Fernandez Soto, J. M., Cepeda, T., and Vano, E. (2004). Patient dose in digital mammography. *Med. Phys.*, 31(9):2471–9.

Clark, B. G. (1980). An efficient implementation of the algorithm "CLEAN". *Astron. Astrophys.*, 89:377–378.

Coleman, N. J., Cowen, A. R., and Parkin, G. J. S. (2000). Effect on mammographic CR image quality of removing the anti-scatter grid. *Radiography*, 6:199–204.

Coleman, T. F. and Li, Y. (1994). On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds. *Mathematical programming*, 67:189–224.

Coleman, T. F. and Li, Y. (1996). An interior, trust region approach for nonlinear minimization subject to bounds. *SIAM Journal on Optimization*, 6:418–45.

Compton, A. H. (1923). A quantum theory of the scattering of X-rays by light elements. *The Physical Review*, 21(5):483–502.

Cooley, J. W. and Tukey, J. W. (1965). An algorithm for the machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301.

Cooper, V. N., Boone, J. M., Seibert, J. A., and Pellot-Barakat, C. J. (2000). An edge spread technique for measurement of the scatter-primary ratio in mammography. *Med. Phys.*, 27(5):845–53.

Cornelis, G., Bellet, A., and VanEygen, B. (1972). Rotational multiple sequence roentgenology of intracranial aneurysm. *Acta Radiol. Diagn.*, 13:74–6.

Cornwell, T. J. (1982). Can CLEAN be improved? VLA Scientific Memorandum No. 141.

Cornwell, T. J. and Evans, K. F. (1985). A simple maximum entropy deconvolution algorithm. *Astron. Astrophys.*, 143:77–83.

Costaridou, L., Sakellaropoulos, P., Stefanoyiannis, A. P., Ungureanu, E., and Panayiotakis, G. (2001). Quantifying image quality at breast periphery vs mammary gland in mammography using wavelet analysis. *British Journal of Radiology*, 74:913–9.

Cowen, A. R., Brettle, D. S., Coleman, N. J., and Parkin, G. J. S. (1992). A preliminary investigation of the imaging performance of photostimulable phosphor computed radiography using a new design of mammographic quality control test object. *British Journal of Radiology*, 65:528–535.

Cowen, A. R., Launders, J. H., Jadav, M., and Brettle, D. S. (1997). Visibility of microcalcifications in computed and screen-film mammography. *Physics in Medicine and Biology*, 42:1533–48.

Cowlinshaw, M. F. (1985). Fundamental requirements for picture presentation. *Proceedings of the Society for Information Display*, 26(2):101–7.

Cranley, K., Gilmore, B. J., Fogarty, G. W. A., Desponds, L., and Sutton, D. (1997). *IPEM Report 78 Catalogue of Diagnostic X-ray Spectra and other Data.* Institute of Physics and Engineering in Medicine, York, UK, cd-rom edition edition.

Csiszàr, I. (1991). Why least squares and maximum entropy? an axiomatic approach to inference for linear inverse problems. *The Annals of Statistics*, 19(4):2032–66.

Dance, D. R. (1988). Diagnostic radiology with x-rays. In Webb, S., editor, *The Physics of Medical Imaging*, Medical Science Series, pages 20–73. IOP Publishing.

Dance, D. R., Skinner, C. L., Young, K. C., Beckett, J. R., and Kotre, C. J. (2000a). Additional factors for the estimation of mean glandular breast dose using the uk mammography dosimetry protocol. *Phys. Med. Biol.*, 45:3225–40.

Dance, D. R., Thilander, A. K., Sandborg, M., Skinner, C. L., Castellano, I. A., and Carlsson, G. A. (2000b). Influence of anode/filter material and tube potential on contrast, signal-to-noise ratio and average absorbed dose in mammography: a Monte Carlo study. *Br J Radiol*, 73(874):1056–67.

De Pierro, A. R. (1995). A modified Expectation Maximization algorithm for penalized likelihood estimation in emission tomography. *IEEE Trans. Med. Imaging*, 14(1):132–7.

De Stefano, A., White, P. R., and Collis, W. B. (2004). Selection of thresholding scheme for image noise reduction on wavelet components using bayesian estimation. *Journal of Mathematical Imaging and Vision*, 21(3):225–33.

Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39(1):1–38.

Dobbins, J. T. and Godfrey, D. J. (2003). Digital x-ray tomosynthesis: current state of the art and clinical potential. *Phys. Med. Biol.*, 48:R65–R106. Topical Review.

Dobbins, J. T. I., Powell, A. O., and Weaver, Y. K. (1987). Matrix inversion tomosynthesis: initial image reconstruction. RSNA 73rd Scientific Assembly (Chicago, IL).

Donoho, D. L. (1992). De-noising via soft thresholding. Technical Report 409, Stanford University, Department of Statistics, Stanford University.

Donoho, D. L. and Johnstone, I. M. (1992). Ideal spatial adaptation via wavelet shrinkage. Technical Report 400, Stanford University, Department of Statistics, Stanford University.

Donoho, D. L., Johnstone, I. M., Hoch, J. C., and Stern, A. S. (1992). Maximum entropy and the nearly black object. *J. R. Statist. Soc. B*, 54(1):41–81.

Dougherty, G. and Kawaf, Z. (2001). The point-spread function revisited: image restoration using 2-D deconvolution. *Radiography*, 7(4):255–262.

Edholm, P. and Quiding, L. (1969). Reduction of linear blurring in tomography. *Radiology*, 92:1115–8.

Edholm, P. and Quiding, L. (1970). Elimination of blur in linear tomography. *Acta Radiologica*, 10:441–7.

Ekestubbe, A., Grondahl, H. G., and Molander, B. (2003). Quality of digital pre-implant tomography: comparison of film-screen images with storage phosphor images at normal and low dose. *Dentomaxillofac Radiol*, 32(5):322–6.

Fandos-Morea, A., Prats-Esteve, M., Tura-Soteras, J. M., and Traveria-Cros, A. (1988). Breast tumours: composition of microcalcifications. *Radiology*, 169:325–7.

Farr, R. F. and Allisy-Roberts, P. J. (1998). *Physics for medical imaging.* W.B. Saunders Company Ltd.

Ferrari, R. J. and Winsor, R. (2005). Digital Radiographic Image Denoising Via Wavelet-Based Hidden Markov Model Estimation. To appear in J Digit Imaging.

Fisher, M. R., Mendelson, E. B., Mintzer, R. A., Ricketti, A. J., and Greenberger, P. A. (1985). Use of linear tomography to confirm the diagnosis of allergic bronchopulmonary aspergillosis. *CHEST*, 87:499–502.

Fletcher, R. and Reeves, C. M. (1964). Function minimization by conjugate gradients. *Computer Journal*, 7:149–54.

Frieden, B. R. (1972). Restoring with maximum likelihood and maximum entropy. *Optical Society of America Journal*, 62:511–8.

Fujita, H., Ueda, K., Morishita, J., Fujikawa, T., Ohtsuka, A., and Sai, T. (1989). Basic imaging properties of a computed radiographic system with photostimulable phosphors. *Med. Phys.*, 16(1):52–9.

Funke, M., Breiter, N., Hermann, K. P., Oestmann, J. W., and Grabbe, E. (1998). Storage phosphor direct magnification mammography in comparison with conventional screen-film mammography–a phantom study. *British Journal of Radiology*, 71(845):528–34.

Ghosh Roy, D. N., Kruger, R. A., Yih, B., and Del Rio, P. (1985). Selective plane removal in limited angle tomographic imaging. *Medical Physics*, 12(1):65–70.

Gill, P. E., Murray, W., and Wright, M. H. (1981). *Practical Optimization.* Academic Press: London.

Goldwasser, B., Cohan, R. H., Dunnick, N. R., Andriani, R. T., Carson 3rd, C. C., and Weinerth, J. L. (1989). Role of linear tomography in evaluation of patients with nephrolithiasis. *Urology*, 33(3):253–6.

Gong, X., Vedula, A. A., and Glick, S. J. (2004). Microcalcification detection using cone-beam ct mammography with a flat-panel imager. *Phys. Med. Biol.*, 49:2183–95.

Gonzalez, R. C. and Woods, R. E. (1992). *Digital image processing.* Addison-Wesley, 2nd edition.

Gonzalez, R. C. and Woods, R. E. (2001). *Digital image processing.* Prentice Hall, 2nd edition.

Government Statistical Service (2004). Breast Screening Programme, England: 2002-03. Crown Copyright for the Department of Health. Statistical Bulletin 2004/06.

Gull, S. F. and Daniell, G. J. (1978). Image reconstruction from incomplete and noisy data. *Nature*, 272:686–90.

Gull, S. F. and Skilling, J. (1999). *MEMSYS5 Users' Manual.* Maximum Entropy Data Consultants Ltd, South Hill, 42 Southgate Street, Bury St. Edmunds, Suffolk, UK, 1.2 edition.

Hansen, P. C. (2002). Deconvolution and regularization with toeplitz matrices. *Numerical Algorithms*, 29:323–78.

Hansson, B., Finnbogason, T., Schuwert, P., and Persliden, J. (1997). Added copper filtration in digital paediatric double-contrast colon examinations: effects on radiation dose and image quality. *European Radialogy*, 7:1117.

Hatcher, D. C. and Aboudara, C. L. (2004). Diagnosis goes digital. *Am J Orthod Dentofacial Orthop*, 125(4):512–5.

Hawkins, W. G. and Leichner, P. K. (1994). An intrinsic 3D Wiener filter for the deconvolution of spatially varying collimator blur. In *Image Processing*, volume 2 of *IEEE International Conference*, pages 163–7, Austin, TX. Med. Center, Nebraska University, Omaha, NE, USA. ICIP-94.

Hay, G. A. and Hughes, D. (1978). *First-year physics for Radiographers.* Baillière Tindall, London, 2nd edition.

Helstrom, C. W. (1967). Image restoration by the method of least squares. *J. Opt. Soc. Am.*, 57(3).

Hermann, G., Janus, C., Schwartz, I. S., Papatestas, A., Hermann, D. G., and Rabinowitz, J. G. (1988). Occult malignant lesions in 114 patients: relationship to age and the presence of microcalcifications. *Radiology*, 169:312–24.

Highnam, R., Brady, M., and English, R. (1999). Detecting film-screen artefacts in mammography using a model-based approach. *IEEE Trans. on Medical Imaging*, 18(10):1016–24.

Hill, R. O. (1996). *Elementary linear algebra.* Harcourt College Publishers, 3rd edition.

Högbom, J. A. (1974). Aperture synthesis with a non-regular distribution of interferometer baselines. *Astron. Astrophys. Suppl.*, 15:417–426.

Horner, K., Lawinski, C. P., and Smith, N. J. (1988). Erbium filtration for dose reduction in dental radiography. *Br J Radiol*, 61(727):609–12.

Hua, Z., Roehrig, H., and Hayworth, M. (1986). Method for improving image quality in digital linear tomography. *Image and Vision Computing*, 4(1):25–8.

Huang, X., Madoc, A. C., and Cheetham, A. D. (2003). Wavelet-based Bayesian estimator for Poisson noise removal from images. In *ICME '03 Proceedings: International Conference on Multimedia and Expo*, volume 1, pages 593–6. Sch. of Inf. Sci. & Eng., Canberra Univ., ACT, Australia.

Hudson, H. M. and Larkin, R. S. (1994). Accelerated image reconstruction using ordered subsets of projection data. *IEEE Trans. Med. Imaging*, 13(4):601–9.

Hurley, S. F. and Kaldor, J. M. (1992). The benefits and risks of mammographic screening for breast cancer. *Epidemiol. Rev.*, 14:101–30.

ICRU (1986). Modulation transfer function of screen- film systems. Technical Report 41, International Commission on Radiation Units and Measurements, Inc.

IRMER (2000). The ionising radiation (medical exposure) regulations 2000. London: HMSO.

Jackson, J. C. and Ward, G. (1981). Surface inspection of steel products using a synthetic aperture microwave technique. *Br. J. of Non-destructive Testing*, 33(8):395–402.

Jain, A. K. (1988). *Fundamentals of Digital Image Processing.* Prentice-Hall information and system sciences series. Prentice-Hall, Englewood Cliffs, NJ.

Jannetta, A. L., Jackson, J. C., Kotre, C. J., Birch, I. P., Robson, K. J., and Padgett, R. (2004). Mammographic image restoration using maximum entropy deconvolution. *Phys. Med. Biol.*, 49(21):4997–5010.

Jansen, M. and Bultheel, A. (1997). Multiple wavelet threshold estimation by generalized cross validation for data with correlated noise. Technical Report TW250, K.U. Leuven.

Jansen, M., Uytterhoeven, G., and Bultheel, A. (1999). Image de-noising by integer wavelet transforms and generalized cross validation. *Medical Physics*, 26:622–30.

Jansson, P. A. (1970). Method for determining the response function of a high-resolution infrared spectrometer. *J. Opt. Soc. Am.*, 60:184–91.

Jaynes, E. T. (1957a). Information theory and statistical mechanics. *Phys. Rev.*, 106:620.

Jaynes, E. T. (1957b). Information theory and statistical mechanics II. *Phys. Rev.*, 108:171.

Jaynes, E. T. (1968). Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics*, 4(3):227–41.

Jaynes, E. T. (1982). On the rationale of maximum entropy methods. *Proceedings of the IEEE*, 70(9):939–52.

Jaynes, E. T. (1986). Monkeys, kangaroos and N. Technical Report 1189, University of Cambridge Physics Dept.

Jones, C. H. (1982). Methods of breast imaging. *Phys. Med. Biol.*, 27:463–99.

Kamel, N. S. and Sim, K. S. (2004). Image signal-to-noise ratio and noise variance estimation using autoregressive model. *Scanning*, 26(6):277–81.

Kamm, J. and Nagy, J. G. (1998). Kronecker product and SVD approximations in image restoration. *Linear Algebra and its Applications*, 284:177–92.

Kamm, J. and Nagy, J. G. (2001). Optimal Kronecker product approximation of block Toeplitz matrices. *SIAM Journal on Matrix Analysis and Applications*, 22(1):155–72.

Keel, W. C. (1991). A simple, photometrically accurate algorithm for deconvolution of optical images. *Publications of the Astronomical Society of the Pacific*, 103:723–729.

Kinahan, P. E., Fessler, J. A., and Karp, J. S. (1997). Statistical image reconstruction in PET with compensation for missing data. *IEEE Tr. Nucl. Sci.*, 44(4):1552–7.

Koh, T. S., Wu, X. Y., Cheong, L. H., and Lim, C. C. (2004). Assessment of perfusion by dynamic contrast-enhanced imaging using a deconvolution approach based on regression and singular value decomposition. *IEEE Trans Med Imaging*, 23(12):1532–42.

Krol, A., Bassano, D. A., Chamberlain, C. C., and Prasad, S. C. (1996). Scatter reduction in mammography with air gap. *Med. Phys.*, 23(7):1263–70.

Kumazaki, T. (1989). Rotational stereo-digital angiography-image processing and clinical usefulness. *Med. Image. Tech. Jpn.*, 7:433–9.

Kumazaki, T. (1991). Development of a new digital angiography system improvement of rotational angiography and three dimensional image display. *Nippon Acta Radiol.*, 51:1068–77.

Lagendijk, R. L., Biemond, J., and Boekee, D. E. (1988). Regularized image restoration with ringing reduction. *IEEE Transactions on acoustics, speech and signal processing*, 36(12):1874–88.

Lantéri, H., Roche, M., Gaucherel, P., and Aime, C. (2002). Ringing reduction in image restoration algorithms using a constraint on the inferior bound of the solution. *Signal Processing*, 82(10):1481–1504.

Lawaczeck, R., Arkadiev, V., Diekmann, F., and Krumrey, M. (2005). Monochromatic x-rays in digital mammography. *Invest. Radiol.*, 40(1):33–9.

Leonard, T. and Hsu, J. S. J. (1999). *Bayesian Methods: An analysis for statisticians and interdisiplinary researchers.* Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Luangjamekorn, N., Williams, N., and Angelopoulos, N. (2005). Digital panoramic radiography image quality: A comparison with film-based panoramic radiography. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 99(3):E25.

Lucy, L. B. (1974). An iterative technique for the rectification of observed distributions. *Astronomical Journal*, 79:745.

MacKay, D. J. C. (2003). *Information theory, inference, and learning algorithms.* Cambridge University Press, Cambridge UK.

Madsen, M. T. (1990). A method for obtaining an approximate Wiener filter. *Med Phys*, 17(1):126–30.

Medical Devices Agency (2001). MDA Evaluation Report mda 01011. further revisions to guidance notes for health authorities and NHS Trusts on mammographic x-ray equipment for breast screening. Department of Health.

Meijering, E., Niessen, W., Weickert, J., and Viergevera, M. (2002). Diffusion-enhanced visualization and quantification of vascular anomalies in three-dimensional rotational angiography: Results of an in-vitro evaluation. *Medical image analysis*, 6(3):215–33.

Meredith, W. J. and Massey, J. B. (1977). *Fundamental physics of radiology.* John Wright and Sons Ltd, Bristol, UK, 1st edition.

Miller, M. I. and Roysam, B. (1991). Bayesian image reconstruction for emission tomography incorporating good's roughness prior on massively parallel processors. *Proc. Natl. Acad. Sci. USA*, 88:32237.

Mol, A. (2004). Imaging methods in periodontology. *Periodontol 2000*, 34:34–38.

Molander, B., Grondahl, H. G., and Ekestubbe, A. (2004). Quality of film-based and digital panoramic radiography. *Dentomaxillofac Radiol*, 33(1):32–6.

Moore, E. H. (1920). On the reciprocal of the general algebraic matrix. *Bull. Amer. Math. Soc.*, 26:394–5.

Moore, W. S. (2002). Dental digital radiography. *Tex. Dent. J.*, 119(5):404–12.

Moran, P. R. (1991). Observations on maximum entropy processing of MR images. *Magn Reson Imaging*, 9(2):213–21.

Mould, R. F. (1995). The early history of X-ray diagnosis with emphasis on the contributions of physics 1895-1915. *Phys. Med. Biol.*, 40:1741–87.

Muntz, E. P., Jafroudi, H., Jennings, R., and Berstein, H. (1985). An approach to specifying a minimum dose system for mammography using multiparameter optimization techniques. *Med. Phys.*, 12:5–12.

Murray, D. and Whyte, A. (2001). Dental Panoramic Tomography: What the general radiologist needs to know. *Clinical Radiology*, 57:1–7.

Nagy, J. G. (1996). Decomposition of block Toeplitz matrices into a sum of Kronecker products with applications in image restoration. Technical Report 96-1, Department of Mathematics, Southern Methodist University.

Narayan, R. and Nityananda, R. (1986). Maximum entropy image restoration in astronomy. *Ann. Rev. Astron. Astrophys.*, 24:127–70.

Niklason, L. T., Christian, B. T., Niklason, L. E., Kopans, D., Castleberry, D. E., Opsahl-Ong, B. H., Landberg, C. E., Slanetz, P. J., Giardino, A. A., Moore, R., Albagli, D., DeJule, M. C., Fitzgerald, P. F., Fobare, D. F., Giambattista, B. W., Kwasnick, R. F., Liu, J., Lubowski, S. J., Possin, G. E., Richotte, J. F., Wei, C. Y., and Wirth, R. F. (1997). Digital tomosynthesis in breast imaging. *Radiology*, 205:399–406.

O'Sullivan, J. A., Blahut, R. E., and Snyder, D. L. (1998). Information-theoretic image formation. *IEEE Transactions on Information Theory*, 44(6):2094–2123.

Penrose, R. (1955). A generalised inverse for matrices. *Proc. Cambridge Phil. Soc.*, 51:406–13.

Phillips, D. L. (1962). A technique for the numerical solution of certain integral equations of the first kind. *Journal of the ACM*, 9(1):84–97.

Pizurica, A., Philips, W., Lemahieu, I., and Acheroy, M. (2003). A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Trans. Med. Imag.*, 22(3):323–31.

Pratt, W. K. (1991). *Digital image processing.* Wiley, New York.

Richardson, W. H. (1972). Bayesian-based iterative method of image restoration. *Optical Society of America Journal*, 62:55–59.

Rowlands, J. A. (2002). The physics of computed radiography. *Physics in Medicine and Biology*, 47:R123–66.

Ruttimann, U. E., Groenhuis, R. A. J., and Webber, R. L. (1984). Restoration of digital multiplane tomosynthesis by a constrained iteration method. *IEEE Transactions on Medical Imaging*, MI-3(3):141–148.

Säbel, M. and Aichinger, H. (1996). Recent developments in breast imaging. *Phys. Med. Biol.*, 41:315–68.

Sahiner, B. and Yagle, A. E. (1993). Limited angle tomography using wavelets. In *Nuclear Science Symposium and Medical Imaging Conference*, volume 3, pages 1912–6. Dept. of Electr. Eng. & Comput., Michigan Univ., Ann Arbor, MI, USA.

Sakellaropoulos, P., Costaridou, L., and Panayiotakis, G. (2003). A wavelet-based adaptive method for mammographic contrast enhancement. *Phys. Med. Biol.*, 48:787–803.

Samei, E., Seibert, J. A., Willis, C. E., Flynn, M. J., Mah, E., and Junck, K. L. (2001). Performance evaluation of computed radiography systems. *Medical Physics*, 28(3):361–71.

Sandborg, M., Carlson, C. A., and Alm-Carlson, G. (1989). Choice of optimal energy spectra in diagnostic radiology: an analysis based on calculated signal-to-noise ratios in clinical, partially absorbing detectors. Technical Report 20, British Institute of Radiology. p.141.

Sandborg, M., Carlsson, C. A., and Carlsson, G. A. (1994). Shaping x-ray spectra with filters in x-ray diagnostics. *Med Biol Eng Comput*, 32(4):384–90.

Schwarz, U. J. (1978). Mathematical-statistical description of the iterative beam removing technique (method CLEAN). *Astron. Astrophys.*, 65:345–56.

Seibert, J. A. and Boone, J. M. (1988). X-ray scatter removal by deconvolution. *Med. Phys.*, 15(4):567–75.

Seibert, J. A., Boone, J. M., Cooper, V. N., and Lindfors, K. K. (2004). Cassette-based digital mammography. *Technol. Cancer Res Treat*, 3(5):413–27.

Sezan, M. I. and Stark, H. (1982). Image restoration by the method of convex projections: Part 2 - Applications and numerical results. *IEEE transactions on medical imaging*, MI-1(2):95–101.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.*, 27:379–423, 623–656.

Shepp, L. A. and Vardi, Y. (1982). Maximum likelihood reconstruction for emision tomography. *IEEE Trans. Medical Imaging*, 1(2):113–22.

Shore, J. E. and Johnson, R. W. (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Transactions on Information Theory*, 26:26–37.

Shrimpton, P. C., Jones, D. G., and Wall, B. F. (1988). The influence of tube filtration and potential on patient dose during x-ray examinations. *Phys. Med. Biol.*, 33(10):1205–12.

Sijbers, J., den Dekker, A. J., Van Audekerke, J., Verhoye, M., and Van Dyck, D. (1998). Estimation of the noise in magnitude MR images. *Magnetic Resonance Imaging*, 16(1):87–90.

Simonetti, G., Cossu, E., Montanaro, M., Caschili, C., and Giuliani, V. (1998). Whats new in mammography? *European Journal of Radiology*, 27(2):S234–S241.

Simpson, D. E., Fleming, J. S., Aldous, A. J., and Daniell, G. J. (1995). Deconvolution of planar scintigrams by maximum entropy. *Phys. Med. Biol.*, 40:153–62.

Skilling, J. (1988). The axioms of maximum entropy. In Erickson, G. J. and Smith, C. R., editors, *Maximum Entropy and Bayesian Methods in Science and Engineering, Volume 1: Foundations*, pages 173–187. Kluwer Academic Publishers.

Skilling, J. (1989). Classic Maximum Entropy. In Skilling, J., editor, *Maximum Entropy and Bayesian Methods*. Kluwer Academic: Dordrecht, The Netherlands.

Skilling, J. and Bryan, R. K. (1984). Maximum entropy image reconstruction - general algorithm. *Mon. Not. R. astr. Soc.*, 211:111–24.

Skilling, J. and Gull, S. F. (1985). Algorithms and applications. In Smith, C. R. and Grandy, W. T., editors, *Maximum Entropy and Bayesian Methods in Inverse Problems*, page 83. D. Reidel Publishing Company.

Snyder, D. L., Schulz, T. J., and O'Sullivan, J. A. (1992). Deblurring subject to nonnegativity constraints. *IEEE Transactions on Signal Processing*, 40(5):1143–50.

Starck, J. L. and Murtagh, F. (1994). Image restoration with noise suppression using the wavelet transform. *Astron. Astrophys.*, 288:342–8.

Starck, J. L. and Pantin, E. (1996). Multiscale maximum entropy images restoration. *Vistas in Astronomy*, 40(4):563–9.

Strickland, R. N. (1996). Wavelet transforms for detecting microcalcifications in mammograms. *IEEE Transactions on medical imaging*, 15(2):218–229.

Sutton, D. G. and Kempi, V. (1992). Constrained least-squares restoration and renogram deconvolution: a comparison by simulation. *Phys. Med. Biol.*, 37(1):53–67.

Tapiovaara, M. J. and Wagner, R. F. (1993). Snr and noise measurements for medical imaging: I. a practical approach based on statistical decision theory. *Phys. Med. Biol.*, 38(1):71–92.

Thomas, G. (1981). A modified version of Van-Cittert's iterative deconvolution procedure. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 29(4):938–9.

Titterington, D. M. (1985). General structure of regularization procedures in image reconstruction. *Astronomy and Astrophysics*, 144(2):381–7.

Van Cittert, P. H. (1931). Zum einfluss der spaltbreite auf die intensitatswerteilung in spektrallinien ii. *Z. Phys.*, 69:298.

Vielva, P., Barreiro, R. B., Hobson, M. P., Martínez-González, E., Lasenby, A. N., Sanz, J. L., and Toffolatti, L. (2001). Combining maximum-entropy and the mexican hat wavelet to reconstruct the microwave sky. *Mon. Not. R. Astron. Soc.*, 328:1–16.

Wagner, R. F., Smith, S. W., Sandrik, J. M., and Lopez, H. (1983). Statistics of speckle in ultrasound b-scans. *IEEE Trans. on Sonics and Ultrasonics*, 30(3):156–63.

Wang, T. C. and Karayiannis, N. B. (1998). Detection of microcalcifications in digital mammograms using wavelets. *IEEE Transactions on medical imaging*, 17(4):498–509.

Webb, S. (1988). The mathematics of image formation and image processing. In Webb, S., editor, *The physics of medical imaging*, Medical science series, chapter 12. IOP Publishing Ltd.

Wenzel, A. (2000). Digital imaging for dental caries. *Dent. Clin. North. Am.*, 44(2):319–38.

White, R. L. (1994). Image restoration using the Damped Richardson-Lucy Iteration. In *ASP Conf. Ser. 61: Astronomical Data Analysis Software and Systems III.*

Wiener, N. (1949). *Extrapolation, interpolation and smoothing of stationary time series.* The MIT Press and John Wiley and Sons, New York, Cambridge, Mass.

Wilks, R. (1987). *Principles of radiological physics.* Churchill-Livingstone, Edinburgh, 2nd edition.

Wong, W. C. K. and Chung, A. C. S. (2004). A nonlinear and non-iterative noise reduction technique for medical images: concept and methods comparison. In *CARS 2004 - Computer Assisted Radiology and Surgery. Proceedings of the 18th International Congress and Exhibition*, volume 1268 of *International Congress Series*, pages 171–6.

Wu, N. (1997). *The Maximum Entropy Method.* Springer Series in Information Sciences. Springer-Verlag Berlin Heidelberg.

Yaffe, M. J. and Rowlands, J. A. (1997). X-ray detectors for digital radiography. *Phys. Med. Biol.*, 42:1–39.

Youla, D. C. and Webb, H. (1982). Image restoration by the method of convex projections: Part 1 - Theory. *IEEE transactions on Medical Imaging*, MI-1(2):81–94.

Young, K. C. and Ramsdale, M. L. (2003). Performance of mammographic equipment in the UK Breast Screening Programme in 2000/2001. NHS Breast Screening Programme Publication 56.

Yuasa, T., Akiba, M., Takeda, T., Kazama, M., Hoshino, A., Watnabe, Y., Hyodo, K., Dilmanian, F. A., Akatsuka, T., and Itai, Y. (1997). Reconstruction method for fluorescent X-ray computed tomography by least-squares method using singular value decomposition. *IEEE Trans. on Nuclear Science*, 44(1):54–62.

Zhou, X. and Gordon, R. (1989). Detection of early breast cancer: an overview and future prospects. *Crit. Rev. Biomed. Eng.*, 17:203–55.